The Dissertation Committee for Hyunbae Park
certifies that this is the approved version of the following dissertation:

The Imprint of the Ionized Intergalactic Medium on the Temperature Anisotropy of the Cosmic Microwave Background and the Mutual-Impact of Reionization and Small-Scale Structure

Committee:

Paul R. Shapiro, Supervisor

Eiichiro Komatsu

Milos Milosavljevic

Steve Finkelstein

Pawan Kumar

The Imprint of the Ionized Intergalactic Medium on the Temperature Anisotropy of the Cosmic Microwave Background and the Mutual-Impact of Reionization and Small-Scale Structure

by

Hyunbae Park, B.S; M.A.

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

THE UNIVERSITY OF TEXAS AT AUSTIN

August 2015

Dedicated to my parents, Changbom Park and Kyung-suk Kang.

Acknowledgments

I would first like to thank my advisor, Prof. Paul R. Shapiro, for insightful advices on my research. Thanks to him, I was able to grow up as a professional astronomer. I also benefited a lot from his international collaboration working on numerous projects using state-of-art simulations of the group. For example, I had a great time working with Prof. Ilian Iliev, Prof. Garrelt Mellema and Prof. Kyungjin Ahn for my 2nd year project. I am happy to a part of this productive collaboration.

Three months of my stay in University of Tokyo as a part of NSF EAPSI 2013 program is an invaluable experience in my life. I thank the NSF for the opportunity. And, I thank the host, Prof. Naoki Yoshida, and his students who helped me stay productive and enjoy my time in Japan.

I received lots of technical help from posdocs of UT cosmology group including, Dr. Yi Mao, Dr. Jun Koda, Dr. Tanja Rindler-Daller, Dr. Jun Zhang, Dr. D'alosio and Dr. Bovile. Dr. Koda's spoon-feeding manual for the GADGET code gave me a huge boost when running my first GADGET simulations. Running the Freshman Research Initiative stream with Dr. D'alosio was a lot of learning experience.

I am indebted by fellow graduate students and other posdocs as well. Dr. Hyo-jeong Kim, Dr. Donghui Jeong, Myongwon Jeon and Mimi Song

gave me a warm welcome when I first came to Austin, Texas and helped me settle here. Dr. Masatoshi Shoji and Chi-ting Chang were great colleagues in the school and great drinking friends outside the school. As senior graduate students, Dr. Hyo-jeong Kim, Myongwon Jeon and Intae Jung gave me lots of insightful advices for my future. I also enjoyed the time I overlapped with Dr. Junhwan Choi, Dr. Sungryong Hong and Dr. Hwi-hyun Kim who arrived here as posdocs recently.

I also had a happy life in Austin thanks to members of the softball team, PAMA. We supported each other not only in baseball fields, but in the battlefields of life here in US as foreigners. I will never forget their kindness. And, I will never forget the moment that our team won the softball tournament of 2015 Korean Spring Festival with me as the captain. Dear members of PAMA, I wish all of you the best of luck in your future.

I should also acknowledge wonderful jobs done by administrative staffs in this department. I was always amazed by how inconvenient my life here would have been without their fantastic work of getting rid of most of the administrative works so that I could focus on my research. Especially, I owe a lot to the former graduate coordinator, Rachel Walker, who had lots of trouble making sure everything about my PhD program is on track.

Finally, words fail to describe how crucial supports from my parents, Changbom Park and Kyung-suk Kang, were for me in completing my PhD.

The Imprint of the Ionized Intergalactic Medium on the Temperature Anisotropy of the Cosmic Microwave Background and the Mutual-Impact of Reionization and Small-Scale Structure

Publication No. _____

Hyunbae Park, Ph.D.
The University of Texas at Austin, 2015

Supervisor: Paul R. Shapiro

Ionized intergalactic medium (IGM) is an important component in cosmic history. After recombination, the universe went though a dark age until the first stars formed. Since the formation of the first stars, the ionized gas, on one hand, played an important role in the history of the universe and, on the other hand, left its imprints on observables that current and future experiments can measure. In this dissertation, we discuss both of each aspects about ionized gas.

First, we discuss the mutual-impact of reionization and the IGM in small-scale structures. While reionization took place preferentially from densest regions of the universe, IGM in average density regions is expected to have

been ionized externally by galaxies formed in denser regions. Until ionized by external radiation, the IGM is expected to have grown numerous small-scale structures. We simulate how the hydrodynamic feedback on the small-scale structures and its impact on recombination. Then, we also discuss our result on how recombination can impact on the global progress of the reionization. Compared to previous works, we improve on the resolution of simulation. Previous studies took into account only the structures that can form in photo-ionized gas down to $10^8\ M_\odot$ in mass. Here, we present a study that resolves halos down to $10^4\ M_\odot$ to account for structures that were able to form before the reionization heats the gas.

Second, we discuss the kinetic Sunyaev-Zel'dovich effect on the Cosmic Microwave Background (CMB) : temperature fluctuations via the Doppler shift induced by the line-of-sight (LOS) component of the momentum of electrons in the ionized IGM. For the EoR contribution to the signal, we calculate the expected signal from simulations of cosmic reionization that, for the first time, includes the effect of "self-regulation" of reionization: star formation in low-mass galaxies ($10^8\ M_\odot \lesssim M_{\rm halo} \lesssim 10^9\ M_\odot$) and minihalos ($10^5\ M_\odot \lesssim M_{\rm halo} \lesssim 10^8\ M_\odot$) is suppressed if these halos form in regions that are already ionized or Lyman-Werner dissociated. For the post-reionization signal, we revisit the currently used model for non-linear transverse momentum power spectrum with a particular emphasis on the connected term that has been neglected in the literature.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Modern cosmology is driven by the Big Bang paradigm. In this paradigm, the universe started at a finite time in the past and was seeded with the structure we see today during the phase of rapid expansion called Inflation. Expansion continued to cool the universe down until all baryons transitioned to a neutral state and the Cosmic Microwave Background was decoupled from the baryons then. A few hundred million years later, gravitational growth of structure led to formation of the first stars re-ionizing the universe, to start what is called the *Epoch of Reionization* (EoR) (for reviews, see [13, 57]). Quasar absorption spectrum measurements tell us that all the intergalactic gas was ionized to a level of 99.9% level within the first billion years of cosmic history. Continuing growth of structure created clusters of galaxies that are the most massive bound structures we see today.

With the advance of technology, we encounter new scientific data from state-of-art experiments and numerical simulations that call for explanations in the Big Bang paradigm. We are now able to observe fainter objects and model the universe with larger dynamic range and more sophisticated prescriptions. With this new data, we are now becoming able to tackle unanswered questions

about details of the EoR that has previously been too far away for observers to see and too non-linear and complex for theorists to model. In this *dissertation*, We shall contribute to the theoretical basis for interpreting EoR observables using analytical methods and numerical simulations.

In Chapter 2, We model the recombination rate of ionized gas during the EoR. Ionizing hydrogen atoms consuming UV photons with $h\nu > 13.6$ eV is one of the key processes of reionization. When an ion recombines not directly to the ground state, but cascades though one or more $n > 1$ states during the process, it ends up converting one ionizing photon into multiple photons, none of which is energetic enough to ionize another hydrogen atom. This effectively destroys the ionizing photon, thereby delaying the ionization process. The rate of this recombination process goes roughly as the square of density and, as the result, it becomes important to understand the clumpiness to model the rate. Here, it is important to resolve down to a small enough scale to necessary to account for all the density structures that contribute to the clumping of gas. With simulations adopting a resolving power that is higher than previous studies, We aim to help, on one hand, estimate the photon budget required to achieve the end of EoR and, on the other hand, more accurately model large-scale ionization features that left imprints on EoR observables including the anisotropies of the Cosmic Microwave Background (CMB), Cosmic Near-Infrared Background (CIB) and high-redshift 21cm emission.

Recent observations are adding more samples of Lyman-α emitters (LAE) and Lyman break galaxies (LBG) at $z \gtrsim 6$ extending the luminosity

function of galaxies to higher redshift and to fainter luminosity. Using the relation between rest-frame optical luminosity and corresponding UV luminosity calibrated at low redshift, the surveys aim to confirm if the number of ionizing photons from galaxies were enough to complete the EoR by $z \approx 6$ [14, 8, 47]. Current observations only see the brightest ones in the population of galaxies that are not enough to have ionized all the hydrogen atoms even once. But, fainter galaxies are expected to account for most of the UV photons produced during that time. When number density of such galaxies is constrained in the future, it will be important to know how many recombinations take place in the gas on average to know precisely how many UV photons are needed to achieve reionization.

Also, we expect that the recombination rate had time and spatial variation. During the pre-ionization phase, gaseous structures would have grown preferentially faster in denser environments. When ionized, the hydrodynamic feedback would disrupt the structures over time. In addition, the universe at large scales ionized in an inhomogeneous manner. Recent simulations find that the progress of the EoR is characterized by growth of HII bubble up to tens of Mpc until they overlap to finish reionization [3, 16, 27, See Figure 1.1 for a schematic description].

Distribution of those HII bubbles is the key feature imprinted on EoR observables that large-scale simulations like the one in Figure 1.1 aim to model. Free electrons in the bubbles scatter CMB photons imprinting their doppler motions on the temperature map (kinetic Sunyaev-Zelovich effect or kSZ ef-

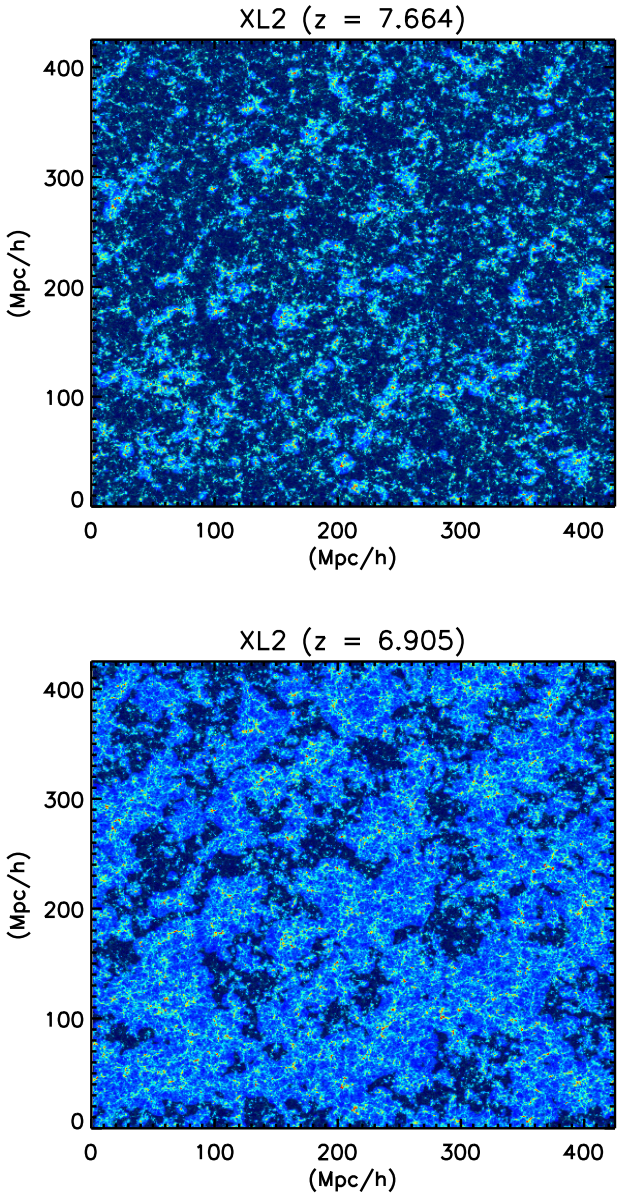

Figure 1.1: (top) A schematic view of the "XL2" simulation of [27] at $z = 7.7$ when the global ionized fraction is 37%. The density field is color-coded such that overdense regions are red while underdense regions are blue. Simultaneously, the ionization field is color-coded such that neutral regions are darkened. (bottom) Similar to the left panel, but for $z = 6.9$ when the global ionized fraction is 81%.

fect). Hydrogen recombination radiation like the Lyman-α emission line will be redshifted into the Cosmic Near-Infrared background at the present epoch. In high-redshift 21cm observations, the HII bubbles will appear as holes empty of signal.

The recombination rate in the bubbles, however, cannot yet be directly calculated from large-scale simulations designed to model EoR observable introduced above. Such simulations are required to be larger than 100 Mpc to capture a meaningful amount of statistics and, therefore, are unable to keep track of structures at sub-Mpc scales that are expected to have contributed significantly to the clumpiness of gas. This unresolved clumpiness is needed to be calibrated in small-scale high-resolution simulations and be provided as a sub-grid prescription. On such small scales, baryonic physics strongly affects the dynamical evolution of gaseous structures. Some of the gas will get ionized immediately when ionizing radiation arrives and some will be shielded by the surrounding gas for a while. When ionizing radiation arrives, increased pressure preferentially causes the ionized parts of the gas to expand. Our goal in Chapter 2 is resolve such situations with fully coupled radative-transfer, hydrodynamics and gravitational dynamics.

Chapter 3 is dedicated to one of the large-scale observables introduced above, the kSZ effect. There are two contributions to the kSZ signal:

1. **Post-reionization contribution**. This is the contribution from redshifts below $z = z_{\rm ov}$, where $z_{\rm ov}$ is the redshift at which reionization is

finished, when individual H II bubbles fully overlap with one another. While the post-reionization contribution depends upon the value of $z_{\rm ov}$, it is not too sensitive to the exact value of $z_{\rm ov}$.

2. **Reionization contribution**. This is the contribution from redshifts above $z = z_{\rm ov}$, where the ionization was patchy and incomplete. This contribution depends not only on $z_{\rm ov}$, but also on the details of the time and spatial variation of inhomogeneous reionization, which are not yet well constrained; thus, one must explore how predictions vary for different models of reionization.

The total signal has been constrained by observations. The South Pole Telescope (SPT) and Atacama Cosmology Telescope (ACT) measure CMB temperature anisotropy at scales beyond the Silk damping scale ($\ell \gtrsim 3000$) where it is possible to distinguish the kSZ signal from the primary signal. The latest constraint is $D^{\rm kSZ}_{\ell=3000} = 2.9 \pm 1.3\ \mu K^2$ (69% CL), where $D_\ell \equiv \ell(\ell+1)C_\ell/2\pi$ [18].

The post-reionization signal depends mostly on the magnitude of density and velocity fluctuations, which can be modeled analytically. We shall improve on this post-reionization model by including a term that has been neglected in the literature. Then, We calculate reionization signals from several EoR scenarios. The sum of reionization and post-reionization signal can be compared to the observational value. If the total signal for an EoR model is out of the observationally-allowed range, one can rule out the EoR model,

thereby constraining the characteristics of early galaxies. This dissertation provides an improved theoretical framework for this ruling-out procedure to be performed when future data comes.

Chapter 2

The Hydrodynamic Feedback of Cosmic Reionization on Small-Scale Structure and Its Impact on Photon Consumption during the Epoch of Reionization

Recombination is a two body reaction. Thus, the rate goes as the square of density with a temperature dependent coefficient in fully ionized gas. Numerical simulations would underestimate the rate if there exists unresolved density structures within resolution elements [23, 63, 32]. To factorize the error due to limited resolution, local clumping factor is often defined as $C \equiv \langle n^2 \rangle / \langle n \rangle^2$ where the bracket denotes the volume average and n is the density of ionized gas[1]. C in a resolution element ($C_{\rm local}$ hereafter) can also be written as $1 + \sigma^2_{<\rm r}$ where $\sigma_{<\rm r}$ is the rms of the unresolved ionized gas density within the resolution element.

Note that C taken from the entire universe ($C_{\rm global}$ hereafter) is often quoted to estimate the number of ionizing photons needed to keep the universe ionized as suggested in [41]. In simulations, $C_{\rm global}$ can be expressed as $1 +$

[1]Strictly speaking, temperature dependence of the recombination coefficient should also impact on the value of C although studies often conveniently ignore it assuming it is a minor effect. See Chapter 2.2 for more refined definition of C taking the temperature dependence into account and quantitative difference made by it.

$\sigma^2_{<\mathrm{r}} + \sigma^2_{>\mathrm{r}}$ where $\sigma_{>\mathrm{r}}$ is the rms of the ionized gas density of all the resolution element in the entire universe. While simulations spanning hundreds of Mpc would capture most of the large-scale variation to that goes into $\sigma^2_{>\mathrm{r}}$, $\sigma^2_{<\mathrm{r}}$ would be so significant that assuming $C_{\mathrm{global}} = 1 + \sigma^2_{>\mathrm{r}}$ would severely underestimate the clumping factor in those simulations. This is often suggested by failures in convergence of C_{global} for increasing resolution (e.g. See Figure 15 of [6]). An error in C_{global} is not easily distinguishable in EOR simulations as its effect is nearly degenerate with the mean efficiency of ionizing sources (i.e. underestimating C_{global} and the source efficiency give similar effects.). The spatial variation of C_{local} across resolution elements, however, may leave an observable impact by affecting the growth of ionized bubble.

Gaseous density structures at sub-Mpc scales are expected to be subject to baryonic physics requiring coupled radiative-transfer and hydrodynamics. A number of numerical works were dedicated to this problem [20, 71, 52, 55, 15, 65, 37, 67] all went for resolving halos down to $\sim 10^8 M_\odot$ adopting $\sim 10^6 M_\odot$ for the mass of the dark matter particle. This is to only resolve halos that can retain gas in the photo-ionization temperature $\sim 10^4 K$. This, however, neglects structures formed during the pre-ionization phase in unheated IGM including *minihalos*. While it is expected that the hydrodynamical feedback from ionization would disrupt such structures formed in low temperature, it should add to the recombination with its initially high clumpiness. Especially, minihalos above $\sim 10^6 M_\odot$ can resist ionizing photons for a significant amount of time ($\gtrsim 10^8$ yr) [63, 32]. [63, 34] were the first to address this problem by

performing fully-coupled radiation-hydro simulations of individual minihalo photo evaporation during the EoR, for this purpose.

To this end, [12] reported dark matter particle mass of $50M_{\odot}$ as the preferable resolution for resolving small-scale structures to obtain reliable $C_{\rm local}$ in a simulation involving structure formation from cosmological initial conditions. Their simulation involved post-processed radiative transfer that allowed them to probe the very initial phase of the gas before it had a chance to respond to the thermal feedback and hydrodynamic consequences. The ray-tracing was done for individual SPH particle, which fixed the mass resolution of the radiative transfer efficiently resolving ionizing fronts in dense regions. In most of their parameter sets of ionizing intensity and arrival time of ionizing background, they reported somewhat high clumping factor ($C_{\rm local} \gtrsim 10$) contrary to other recent works [55, 15, 65, 37, 67] which give $C_{\rm local} \lesssim 10$. This implies that the recombination rate may have been underestimated at least in the early phase of ionization in previous works. My goal is to adopt a resolution and a methodology similar to those suggested by [12] while keeping track of the hydrodynamic evolution of the gas fully coupled to the radiative-transfer to model $C_{\rm local}$.

2.1 Methodology

2.1.1 Gravity, Hydrodynamics, & Chemistry

In order to investigate the problem with a fixed mass-resolution, we adopt the smoothed particle hydrodynamics (SPH) code GADGET-3 [70, 68]

Table 2.1: Simulation Parameter & Results

label	Box size (kpc/h)	# of ptls	z_i	J_{21} (Γ_{-12})	Shielding	Dynamics (cm^{-3})	$C_r^{\mathrm{peak}\,a}$	$N_{\mathrm{rec,add}}{}^{b}$
S_I0_z10	**100**	$\mathbf{2 \times 128^3}$	10	1 (9.2)	on	on	16.7	-
M_I0_z10_NS	200	2×256^3	10	1 (9.2)	**off**	on	-	0.57
M_I0_z10_ND	200	2×256^3	10	1 (9.2)	on	**off**	26.0	-
M_I0_z10	200	2×256^3	10	1 (9.2)	on	on	21.0	0.30
M_I-0.5_z10	200	2×256^3	10	**0.3 (2.8)**	on	on	12.7	0.26
M_I-1_z10	200	2×256^3	10	**0.1 (0.92)**	on	on	7.5	0.21
M_I0_z9	200	2×256^3	**9**	1 (9.2)	on	on	28.1	0.35
M_I0_z8	200	2×256^3	**8**	1 (9.2)	on	on	37.5	0.41
L_I0_z10	**400**	$\mathbf{2 \times 512^3}$	10	1 (9.2)	on	on	21.4	-

[a] Because C_r in M_I0_z10_NS has a monotonic behavior, the peak value cannot be defined for that case. M_I0_z10_ND also has the same problem, but we list its converging value.

[b] S_I0_z10, M_I0_z10_ND and L_I0_z10 is not been run to $\Delta t = 150$ Myr that $N_{\mathrm{rec,add}}$ is measured.

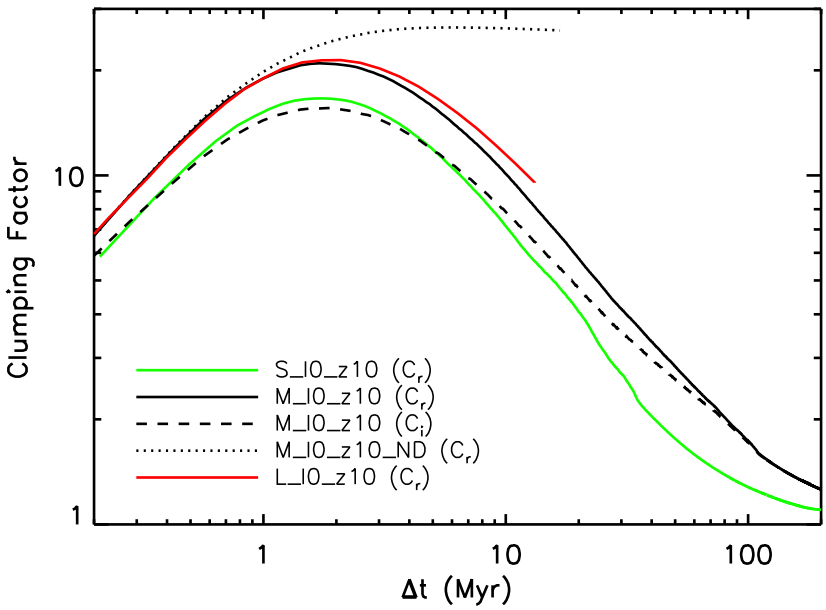

Figure 2.1: Clumping factor minus one plotted as a function of time for several different cases. Solid lines compare C_r from simulations with different box sizes, S_I0_z10(blue), M_I0_z10(black) and L_I0_z10(red). The dotted black line denotes C_r from M_I0_z10_NS to verify the impact of shielding. Finally, the black dashed line is C_r from M_I0_z10.

with non-equilibrium chemistry of 14 primordial species (e^-, H, H^+, H^-, He, He^+, He^{++}, H^2, H_2^+, D, D^+, HD, HD^+, HD^-) as described b.b.y [76, 75] with updated cooling rates for H^2 and HD from [17]. Throughout this paper, the mass resolution is 9.3 $M_\odot$ for baryonic particles and 51 $M_\odot$ for dark matter particles that corresponds to having 256^3 particles for each species in a cubic volume of $(200\ h^{-1}\ \mathrm{kpc})^3$.

The initial condition for $z = 99$ is created using MUSIC [22]. While we mainly use 200 h^{-1} kpc for the size of the box for our analysis, we also use 100 h^{-1} kpc and 400 h^{-1} kpc in order to test the result for different box sizes. We first evolve the initial condition down to $z = 19$ without any radiation background. After $z = 19$, we suppress formation of molecular hydrogen by turning on a uniform Lyman-Werner (LW) background. The spectrum of the LW background is set by black-body temperature of T_{bb} =100,000 K truncated above $\nu = 13.6\ \mathrm{eV}/h_p$ and the normalization set by $J_{21} = 100$ where J_{21} is the intensity at $\nu = 13.6\ \mathrm{eV}/h_p$ in the unit of 10^{-21} erg s^{-1} Hz^{-1} sr^{-1}. Radiative cooling by H_2 molecules leading to baryonic collapse inside the minihalo is strongly prohibited for $J_{21} > 10$ [25]. While the adopted value of $J_{21} = 100$ for the LW background is rather fiducial, we note this represents our target problem, that of a minihalo which has been deactivated in star formation (SF) throughout its history.

With star-formation suppressed, our main sample volume with 200 h^{-1} kpc on a size is evolved down to $z = 7$ and snapshots are saved for $z = 10, 9, 8$ and 7 to be used as initial conditions for the runs in which an external ioniz-

ing background radiation (EIBR) is turned on at these redshifts. Two other sample volumes with 400 h^{-1} kpc and 100 h^{-1} kpc on a size are evolved down to $z = 10$ in the same way.

2.1.2 Algorithm for External Ionizing Background Radiation (EIBR)

We adopt a uniform and isotropic radiation background that is shielded for each particle by the neighboring particles within a certain distance, l_s. Each of the neighboring particles is assigned to the closest one of $\pm x, \pm y$ and $\pm$ z directions to calculate the average column densities of neutral hydrogen for the six directions that are then used to calculate the attenuation from these directions. In Figure 2.2, we show a schematic description for how neighboring particles are assigned to each directions from the target particle. In the upper panel, the target particle is located at left end of the dense clump and will not be shielded from $-x$ direction. On the contrary, the target particle in the lower panel will be shielded from $\pm x, y$ directions ($\pm z$ directions are not shown).

For each neighboring particle shielding the target particle, we add $f_{\rm HI}(m_{\rm gas}/m_p)/(4\pi d_{\rm sh}^2/6)$ to HI column density for the direction that the particle is assigned, where $m_{\rm gas}$, m_p, $f_{\rm HI}$ and $d_{\rm sh}$ are the mass of gas particle, the mass of proton, the number fraction of HI to the number of nucleons and the distance from the shielding particle to the shielded particles, respectively. For the helium, we assume the neutral fraction of it follows that of hydrogen and it is only singly ionized during EOR, which is a reasonable assumption for soft

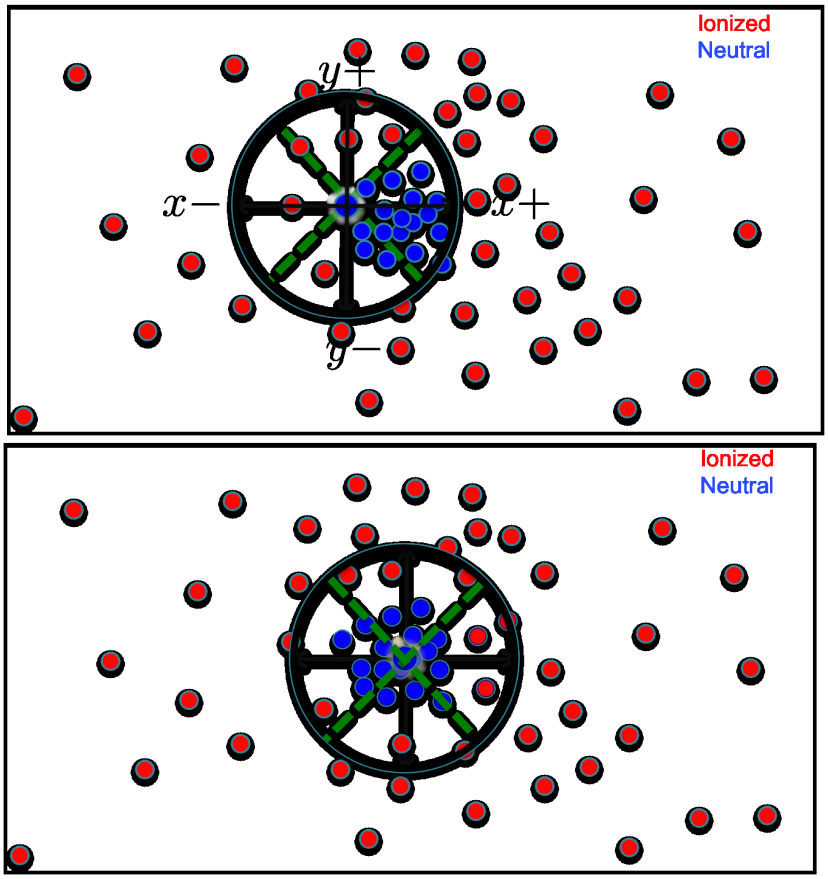

Figure 2.2: (upper) A schematic description for the shielding algorithm used in this work. Blue circles denote neutral SPH particles self-shielding from EIBR where as red circles denote ionized SPH particles. The blue circle at the center of the black circle represents the target particle that we shall calculate the optical depth to EIBR. In this panel a particle at the outer edge of the clump is chosen as the target. The black circle represents the range within which neighboring particles are allowed to shield target particle. Arrows denote $\pm x, y$ directions on the xy-plane that we calculate the optical depth separately. The green dashed lines represent the boundary for each direction. (lower) Same as the left, but the target particle is the one located at the center of the clump.

UV spectrum that is responsible for the reionization in most cases [10]. This algorithm is implemented into the GADGET-3 code to be directly coupled with gravity, hydrodynamics and chemistry solver. We shall call this code as GADGET-RT in this work. We test how accurately this code can keep track of photo-evaporation of a minihalo in Appendix **??**.

2.1.3 Simulations

We simulate the EIBR using output snapshots from Chapter 2.1.1 as initial conditions. In the left seven columns of Table 2.1, we list the name and the parameters for each run. For all the cases, the spectrum of EIBR is given by $T_{\rm bb}$ = 100,000 K with its intensity set by J_{21} = 1, 0.3 or 0.1. Note that this is similar to how we set the LW background in Chapter 2.1.1 except that we do *not* truncate the spectrum above ν =13.6 eV/h_p. We adopt M_I0_z10 that uses J_{21} = 1 and z_i = 10 as the standard run and create other cases by changing one of its characteristics to explore the dependency of the result on the characteristics. The deviating characteristic of each run from the standard run is written in boldface in the table.

For S_I0_z10 and L_I0_z10, we use initial conditions generated from different initializations from different sizes of 100 h^{-1} kpc and 200 h^{-1} kpc and 400 h^{-1} kpc, respectively, to check the convergence of our result on the box size. For M_I0_z8 and M_I0_z9, we change z_i from 10 to 8 and 9, respectively, to investigate the dependence of the result on the timing of reionization. We also study the dependence of J_{21} by changing it to 0.3 (M_I-0.5_z10) and 0.1 (M_I-

1_z10). Lastly, we turn off the shielding algorithm introduced in Chapter 2.1.1 for M_I0_z10_NS and the dynamics of particles (i.e. freeze particle positions as in post-processed radiative transfer) for M_I0_z10_ND.

2.2 Clumping Factor : Definition and ways to calculate

A difference in ionization rate and recombination rate of hydrogen leads to change in the number density of ionized hydrogen:

$$dn_{\rm HII}/dt = I - R. \tag{2.1}$$

The recombination rate of hydrogen is given by

$$R \equiv \alpha_{\rm B}(T) n_e n_{\rm HII}, \tag{2.2}$$

where $\alpha_{\rm B} = 2.6 \times 10^{-13} \times (T/10^4 K)^{-0.7} {\rm s}^{-1} {\rm cm}^3$ is the case B recombination coefficient, T is the temperature of the gas and $n_{\rm X}$ denotes number density of a species X. And, the ionization rate is

$$I \equiv n_{\rm HI} \int d\Omega \int d\nu \sigma(\nu) \frac{J_\gamma(\hat{\Omega}, \nu)}{h\nu}, \tag{2.3}$$

where J_γ is intensity of ionizing radiation. In numerical calculations, the mean recombination rate inside a resolution element is given using the mean physical quantities for the resolution element:

$$\bar{R} = \alpha_{\rm B}(\bar{T}) \bar{n}_e \bar{n}_{\rm HII} \tag{2.4}$$

Here, $\bar{R}$, $\bar{n}_e$ and $\bar{n}_{\rm HII}$ are given by volume average, $\langle \rangle_V$, and the average temperature is given by

$$\bar{T} = (m_p/k_B) \left\langle (\gamma - 1) u \right\rangle_M \left\langle \mu^{-1} \right\rangle_M^{-1}, \tag{2.5}$$

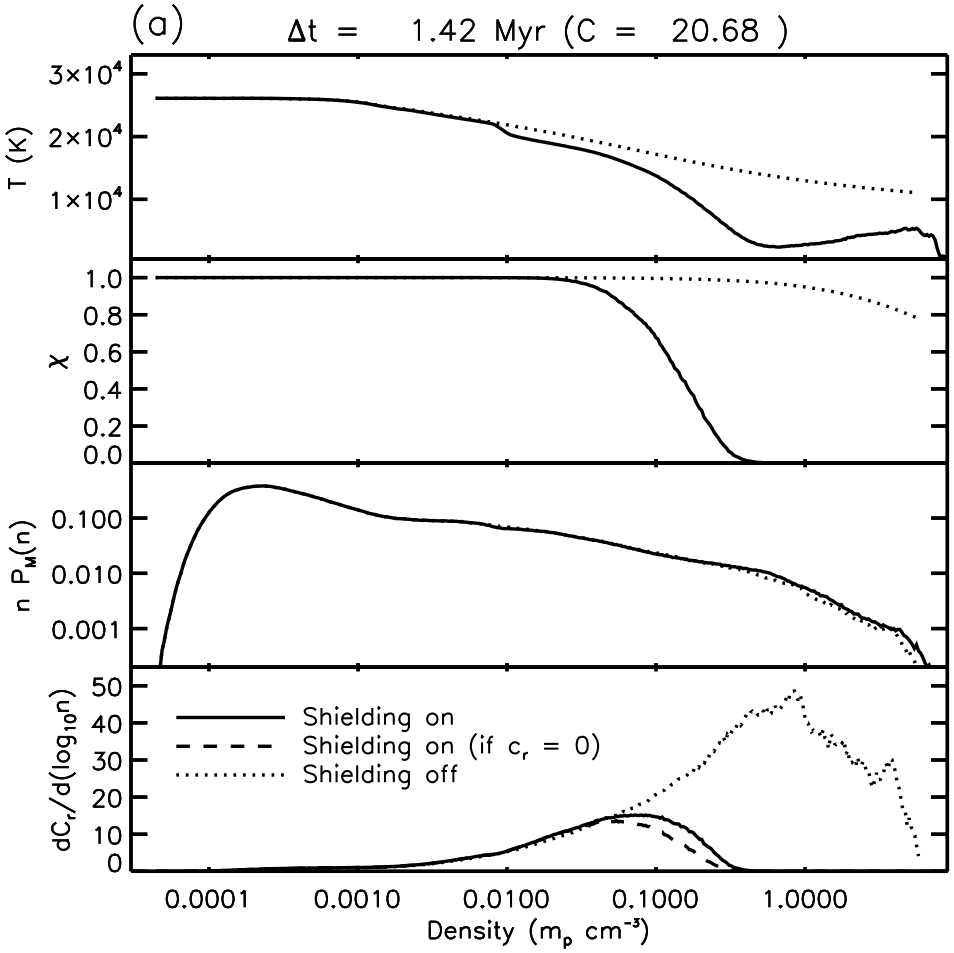

Figure 2.3: Mean gas temperature (T, top panel), mean ionized fraction (χ, upper-middle panel), probability density function (P_M, lower-middle panel) and clumping factor contribution ($dC_r/d\log_{10} n$, bottom panel) as functions of density (n) at $\Delta t = 1.42$ Myr. Solid and dotted lines compare the results from M_I0_z10 and M_I0_z10_NS, respectively. The dashed lines in the bottom panel describes the case that we set $c_r = 1$ in $dC_r/d\log_{10} n$.

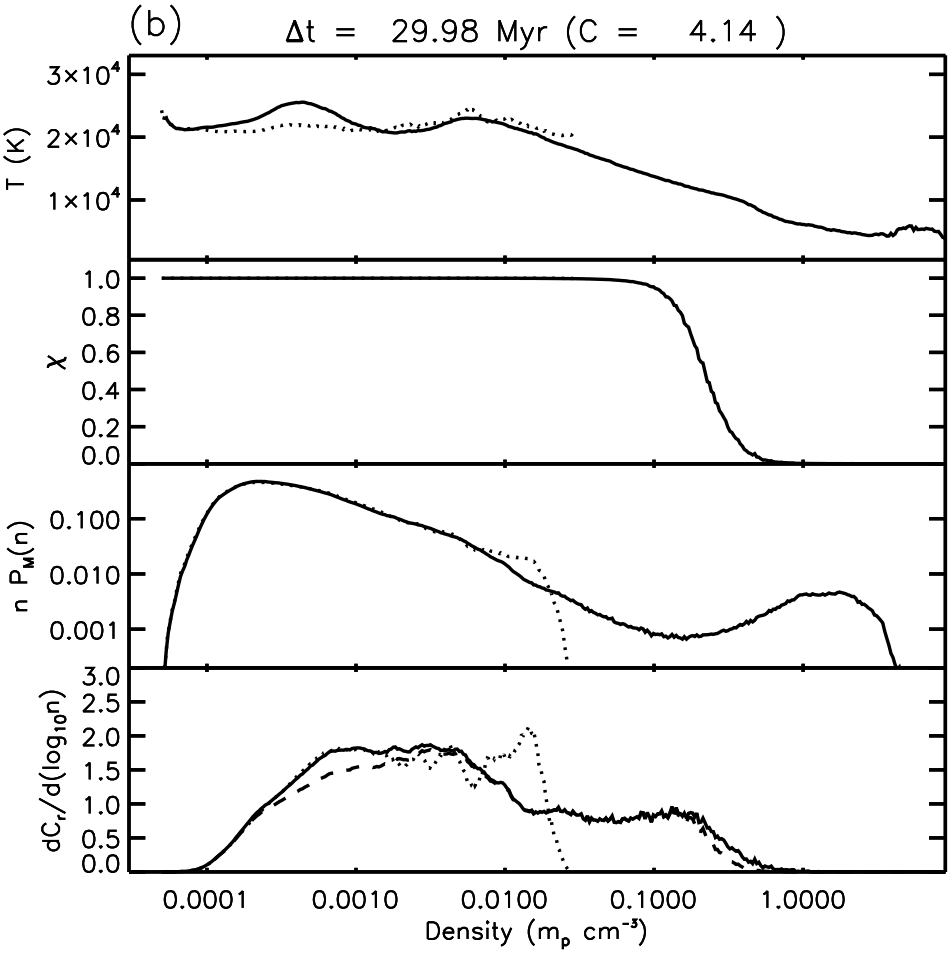

Figure 2.4: Same as Figure 2.3, but for $\Delta t = 37.12$ Myr.

where u is the specific internal energy, μ is mean molecular weight and $\langle\rangle_M$ denotes the mass average.

During the cosmic reionization, it is often assumed that $n_e = (1+Y)n_{\rm HII}$ where $Y = 0.08$ corresponds to the case that helium is singly ionized when hydrogen is ionized and T is constant at $20,000$ K or a similar value. This leads to a clumping factor defined as :

$$C_{\rm i} \equiv \frac{\langle n_{\rm HII}^2 \rangle_V}{\langle n_{\rm HII} \rangle_V^2}. \tag{2.6}$$

Then, the recombination rate is given by

$$R = C_{\rm i}\ \alpha_{\rm B}(T)(1+Y)\bar{\chi}^2\bar{n}_{\rm H}^2, \tag{2.7}$$

where $\chi \equiv n_{\rm HII}/n_{\rm H}$ is the ionized fraction of hydrogen.

In this work, T, n_e and $n_{\rm HII}$ are computed in the simulation and, therefore, we define $C_{\rm r}$ in the following way to more accurately describe the recombination rate :

$$C_{\rm r} \equiv \frac{\langle \alpha_{\rm B}(T) n_e n_{\rm HII} \rangle_V}{\langle n_{\rm HII} \rangle_V \langle n_{\rm e} \rangle_V \alpha_{\rm B}(\bar{T})}. \tag{2.8}$$

In this case,

$$R = C_{\rm r}\ \alpha_{\rm B}(\bar{T})\bar{n}_{\rm HII}\bar{n}_e. \tag{2.9}$$

For defining the SPH smoothed field of a physical quantity $X(\mathbf{r})$ from particle values of X, we adopt a standard method in SPH:

$$X(\mathbf{r}) = \sum_i \frac{X_i}{n_i} W(\mathbf{r} - \mathbf{r}_i; h_i), \tag{2.10}$$

where the subscript i denotes the ith SPH particle in the simulation, $n \equiv \rho/m_p$ is density in the unit of proton mass m_p, $\mathbf{r}$ is location, W is the kernel and h_i is adaptive kernel size given by the distance to the 32nd nearest neighbor from the particle. Then, the volume average of this quantity over the simulation volume $V_{\rm sim}$ is given by

$$\langle X \rangle_V = \frac{1}{V_{\rm sim}} \sum_i \frac{X_i}{n_i} \int_V W(\mathbf{r} - \mathbf{r}_i; h_i) d^3 r. \tag{2.11}$$

By definition, the volume integral of the kernel in the above should give unity, thereby leading to

$$\langle X \rangle_V = \frac{1}{V_{\rm sim}} \sum_i X_i n_i^{-1}. \tag{2.12}$$

This way one can come up with summations that give $C_{\rm i}$ and $C_{\rm r}$ from SPH simulations:

$$C_{\rm i} = \bar{n}^{-1} N_{\rm ptl} \left(\frac{\Sigma_i n_i \chi_i^2}{(\Sigma_i \chi_i)^2} \right) \tag{2.13}$$

$$C_{\rm r} = \bar{n}^{-1} N_{\rm ptl} \left(\frac{\Sigma_i f_{e,i} f_{\rm HII,i} \alpha_{\rm B}(T_i) n_i}{(\Sigma_i f_{e,i})(\Sigma_i f_{\rm HII,i}) \alpha_{\rm B}(\bar{T})} \right). \tag{2.14}$$

Here, $N_{\rm ptl}$ is the number of SPH particles, $f_X \equiv n_X/n$ is the number density of a species X divided by n. For computing $\bar{T}$, the mass average in Eq. 2.5 can be done by simply averaging over the particles.

In this paper, we attempt to understand the clumping factor contribution from gas at the given density, n, by expressing the clumping factor as an integral over n of a combination of relevant physical quantities (χ, f_e,

$f_{\rm HII}$ and T) averaged at a given n. Equation (2.13) and Equation (2.14) can be rewritten into integral form using the mass-weighted probability density function for SPH particles, $P_M(n') = dN_{\rm ptl}/dn'^2$:

$$C_{\rm i} = \frac{\int dn' P_M(n') \chi(n')^2 c_{\rm i}(n') n'}{\bar{n} \langle \chi \rangle_M^2} \tag{2.15}$$

$$C_{\rm r} = \frac{\int dn' P_M(n') \chi(n') f_e(n') \alpha_{\rm B}(T(n')) c_{\rm r}(n') n'}{\bar{n} \langle \chi \rangle_M \langle f_e \rangle_M \alpha_{\rm B}(\bar{T})}. \tag{2.16}$$

Here, the masse average of a quantity X simply come from $N_{\rm ptl}^{-1} \sum_i X_i$. We bin all the SPH particles to 400 density bins uniformly spaced in the log-space between the maximum and the minimum densities of SPH particles in the simulation. Also,

$$c_{\rm i}(n') \equiv \frac{\langle \chi^2 \rangle_{n=n'}}{\bar{\chi}(n')^2};\ c_{\rm r}(n') \equiv \frac{\langle \chi f_e \alpha_{\rm B}(T) \rangle_{n=n'}}{\bar{\chi}(n') \bar{f}_e(n') \alpha_{\rm B}(\bar{T}(n'))}$$

where the bracket $\langle \rangle_{n=n'}$ denotes a mass average within the bin for $n = n'$ that we calculate other physical quantities. These quantify the effect of variation in physical quantities at a given density and should be included in the integral form in order to precisely recover C_i and C_r calculated from Equation (2.13) and Equation (2.14), respectively.

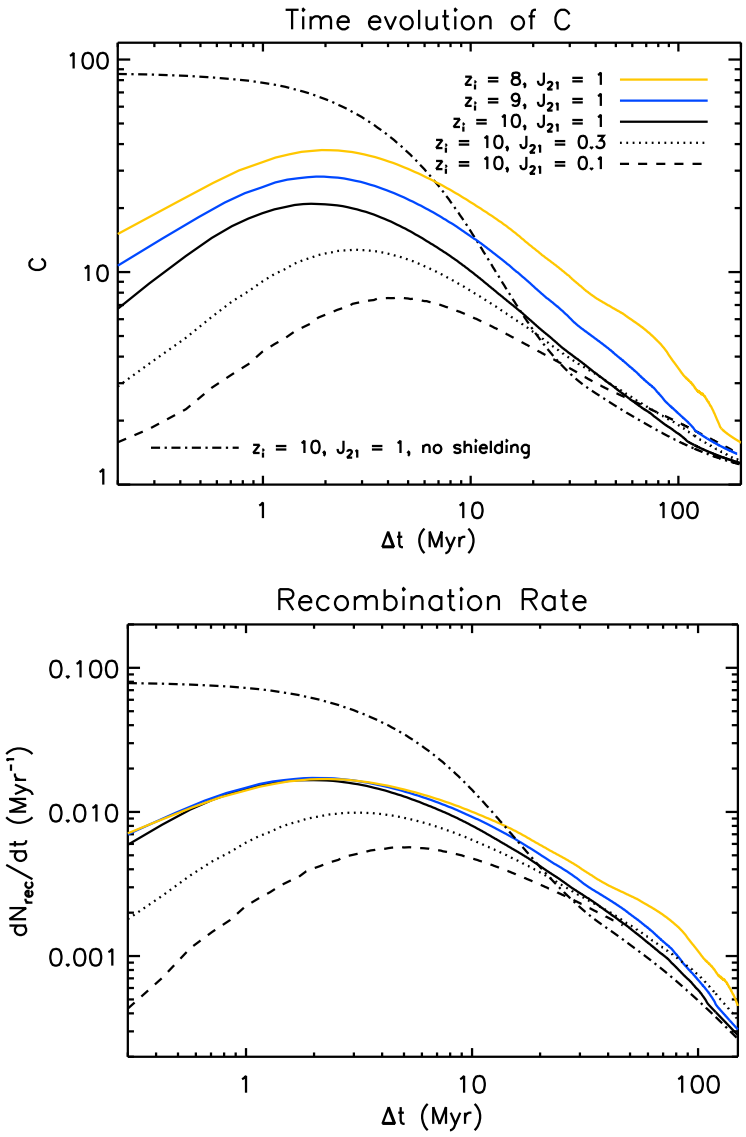

Figure 2.5: (upper) C_r as a function of time for M_I0_z10 (black solid), M_I-0.5_z10 (black dotted), M_I-1_z10 (black dashed), M_I0_z9 (blue solid), M_I0_z8 (yellow solid) and M_I0_z10_NS (black dot-dashed). (lower) Recombination rate as a function of time for same runs plotted in the upper panel.

2.3 Results

2.3.1 Understanding Clumping Factor using density pdf and ionization profile

This section introduces our approach of understanding the time evolution of clumping factor using that of P_M and χ and T as the functions of n. Figure 2.3 and Figure 2.4 show the key terms in the integrals of Equation 2.16 (T, χ, P_M) and the integrand, $dC_r/d\log(n)$, as the functions of n for the standard run (M_I0_z10) and no-shielding case (M_I0_z10_NS) at $\Delta t = 1.42$ Myr and 30 Myr. $dC_r/d\log(n)$ in the bottom panel of Figure 2.3 and Figure 2.4 are plotted in a way that the area the curve gives the total C_r. Additionally, $dC_r/d\log(n)$ with setting $c_r = 1$ is plotted in the dashed line.

The temperature (top panel) shows an anti-correlation with the density, which is expected from the fact that higher density gas has a low equilibrium temperature. Since the Case B recombination coefficient α_B goes as $T^{-0.7}$, this results in a boost of the recombination rate in high density gas. According to bottom panel of Figure 2.3, most of contribution to the clumping factor comes from $0.1 \lesssim n \lesssim 1$ cm^{-3}, which is the high-density end of ionized gas. This explains the result that C_r being higher than C_i, which has been often adopted as the clumping factor. Hereafter, we refer to C_r when discussing the clumping factor unless otherwise noted. From $n \approx 0.03$ cm^{-3} to $n \approx 0.3$ cm^{-3}, χ makes a transition from unity to zero. Within this range, $c_r > 1$ because χ

[2]The volumed weighted gas density pdf of **(author?)** [46] is related to the our mass-weighted pdf by $\Delta P_V(\Delta)d\Delta = P_M(n)dn$, where $\Delta = n/\bar{n}$ is the normalized density.

and f_e highly correlates with each other. The dashed line in the bottom panel of Figure 2.3 suggests that accounting for c_r adds about 10% to the clumping factor compared to the case that we set $c_r = 1$.

The density pdf in Figure 2.3 is close to the initial condition and shows $P_M \propto n^{-1.5}$ for $0.001 \lesssim n \lesssim 1$ cm^{-3} that is expected in isothermal collapse scenario. This is also equivalent with $P_V \propto \Delta^{-2.5}$ reported by [46] using hydrodynamic simulation without radiation. Thus, $dC_r/d\log(n)$ goes as $n^{0.5}$ at densities that the gas is fully ionized ($\chi = 1$). As χ drops from $n \gtrsim 0.03$ cm^{-3}, $dC_r/d\log(n)$ is suppressed. Notably, $dC_r/d\log(n)$ in the no-shielding case (M_I0_z10_NS) continues to rise until P_M falls off and gives a much larger clumping factor.

2.3.2 Time Evolution of Clumping Factor

In Figure 2.1 and Figure 2.5, we show the clumping factor, C_r, as a function of the time elapsed from the turn-on of the EIBR (Δt) for the standard case (M_I0_z10) compared with other cases. Except for the ones without shielding (M_I0_z10_NS) and without dynamics (M_I0_z10_ND), we observe a characteristic behavior in the evolution of the clumping factor. Initially, the clumping factor rises until it peaks at $\Delta t \sim 1$ Myr. Then, it starts to fall eventually converging to unity at late times. After $\Delta t \sim 100$ Myr, the clumping factor falls close to or below 2 in all cases. To understand this dual phase evolution of the clumping factor, we plot the projected HI column density map of M_I0_z10 run at four different times ($\Delta t = 0.14$ Myr, 1.4 Myr, 7.1

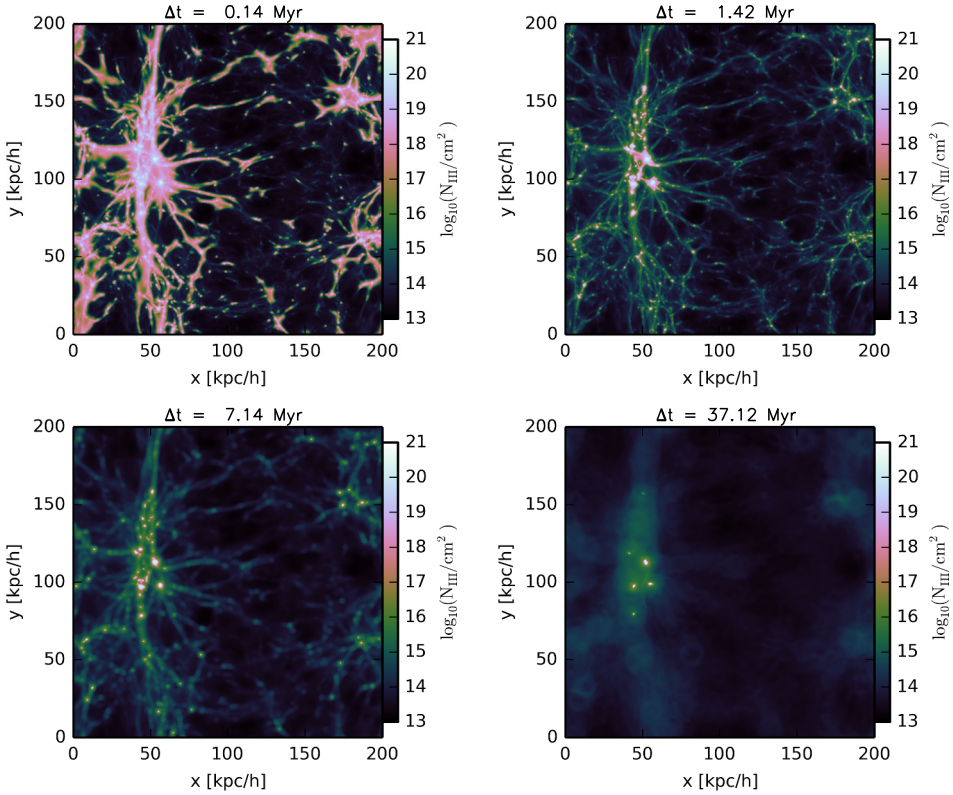

Figure 2.6: Projected neutral column density of M_I0_z10 at $\Delta t = 0.14$ Myr (top left), 1.4 Myr (top right), 7.1 Myr (bottom left) and 37 Myr (bottom right).

Myr & 37 Myr) in Figure 2.6 to give a pictorial description of the ionization status. And, we plot the profiles of ionization (χ) and density pdf (P_M) for M_I0_z10 at those times in Figure 2.7 to give a quantitative description of the time evolution.

The top left panel of Figure 2.6 shows an early time ($\Delta t = 0.14$ Myr) in the simulation. Bright regions with $N_{\rm HI} \gtrsim 10^{19}$ cm^{-2} describe dense structures with self-shielding gas. The top panel of Figure 2.7 shows where the ionized fraction profile, $\bar{\chi}(n)$, transitions to zero at $n \sim 0.03$ cm^{-3}. Thus, the bright regions represent regions with $n \gtrsim 0.03$ cm^{-3} at $\Delta t = 0.14$ Myr. Those neutral regions shrink rapidly in the early time until $\Delta t = 1.4$ Myr. The transition density of χ also shifts to a higher density ($n \sim 0.2$ cm^{-3}) at the same time. We define the density that $\chi = 0.5$ as the critical density of ionization, $n_{\rm crit}$ and plot in Figure 2.11 (For the standard case M_I0_z10, see the black solid line). The density pdf, $P_M(n)$, in the middle panel of Figure 2.7, does not evolve significantly during this time indicating that hydrodynamical feedback is yet to come into the play. All these suggest that ionization fronts are in the supersonic r-type phase. Low density regions are ionized almost immediate after EIBR turns on and ionization fronts propagate into denser regions while the gas is yet to respond to photo-heating. The clumping factor grows during this since as denser and denser gas is ionized and is allowed contribute to the clumping factor. The bottom panel of Figure 2.7 the evolution in dC_r/dn in Equation (2.16) from $\Delta t = 0.14$ Myr to $\Delta t = 1.4$ Myr is mostly explained by that of $\bar{\chi}(n)$: As the break in $\bar{\chi}(n)$ moves from 0.03 cm^{-3} to 0.2 cm^{-3}, gas

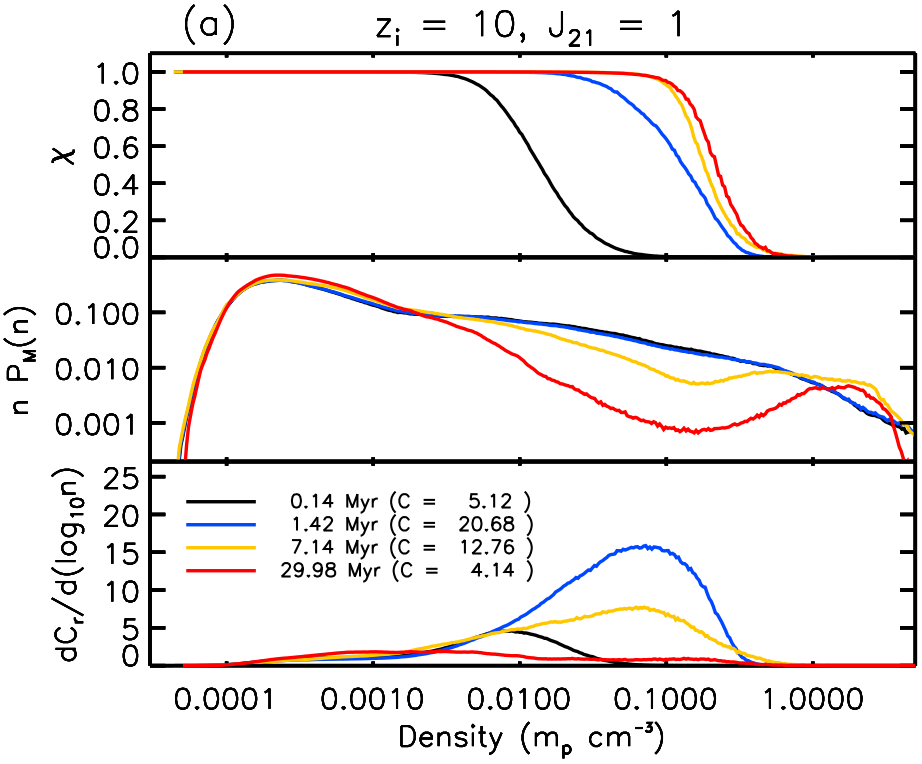

Figure 2.7: Mean ionized fraction (χ), probability density function (P_M) and clumping factor contribution ($dC_r/d\log_{10} n$) as functions of n for M_I0_z10 in top, upper middle, lower middle and bottom panels, respectively. Lines with black, blue, yellow and red color describe snapshots of simulations at four different times Δt = 0.14 Myr, 1.4 Myr, 7.1 Myr and 30 Myr.

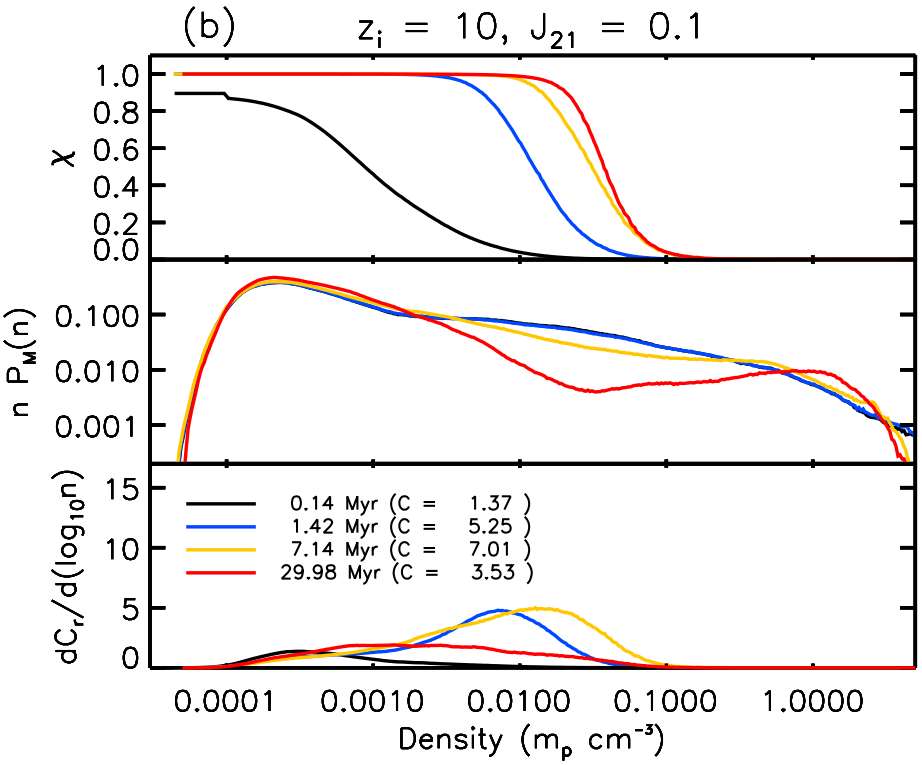

Figure 2.8: Same as Figure 2.7, but for M_I-1_z10.

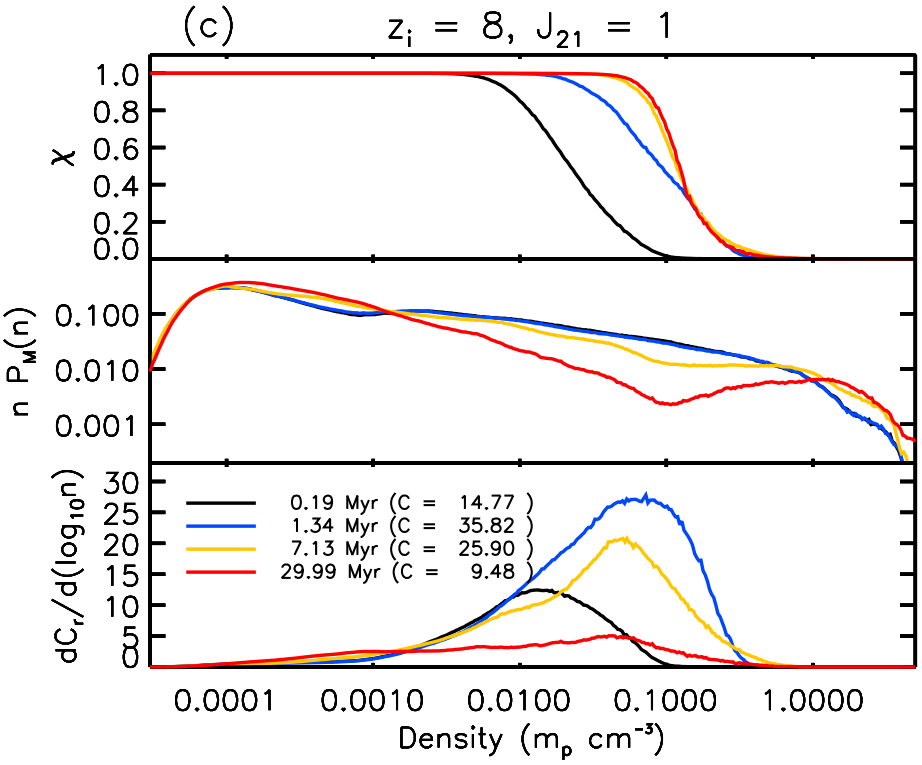

Figure 2.9: Same as Figure 2.7, but for M_I0_z8. And, lines with black, blue, yellow and red color describe snapshots of simulations at four different times $\Delta t = 0.19$ Myr, 1.3 Myr, 7.1 Myr and 30 Myr.

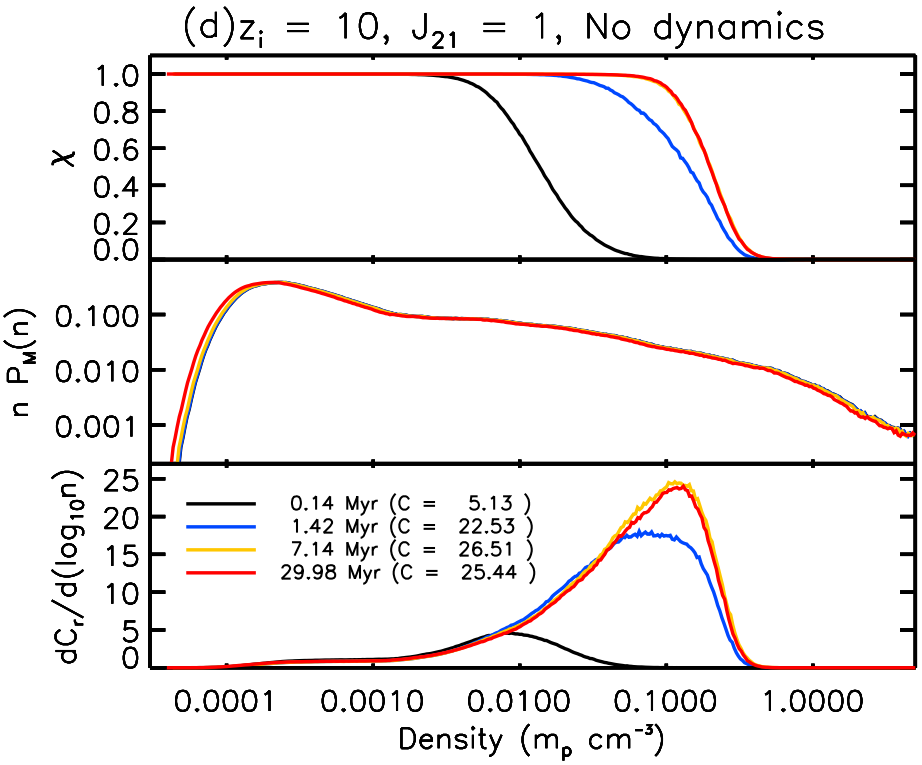

Figure 2.10: Same as Figure 2.7, but for M_I0_z10_ND.

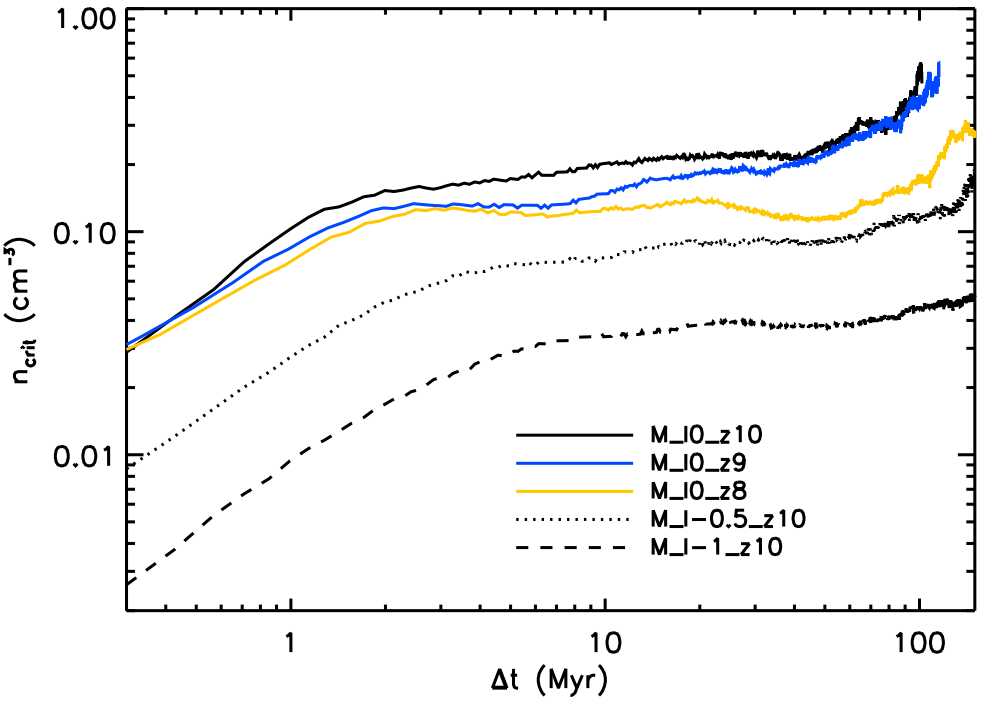

Figure 2.11: Critical density of ionization ($n_{\rm crit}$) as a function of time from the turn-on of EIBR (Δt).

between those densities newly contributes to C_r at $\Delta t = 1.4$ Myr and, hence, the clumping factor becomes larger. This phase lasts until $\Delta t \sim 2$ Myr when ionization fronts slow down and hydrodynamical feedback begins to expand ionized gas.

Later ($\Delta t \gtrsim 2$ Myr), ionization fronts are trapped on surfaces of neutral clumps in minihalos as in the *d*-type phase. The HI density maps show filamentary structures are expanding and being diluted with the background over time. At $\Delta t = 37$ Myr, most of the structures have been puffed up and a few minihalos resisting their complete ionization. $\bar{\chi}(n)$ stops evolving forming a step function at $n_{\rm crit} \approx 0.2$ cm^{-3} up to this point. $n_{\rm crit}$ rises further after $\Delta t \approx 40$ Myr when ionization fronts reach centers of minihalos to finish the last remaining neutral gas. As the hydrodynamical feedback causes dense ionized gas to expend to lower densities, P_M is suppressed at 0.003 cm$^{-3} < n < n_{\rm crit}$ and enhanced at $n < 0.003$ cm^{-3}. This leads to a decay of C_r since high-density ionized gas is the main contributor of C_r.

Density and ionization histories of individual particles support this picture as well. We sample ten SPH particles with difference initial densities ($n_{\rm init}$) and show how they evolve over time in both density and ionized fraction in Figure 2.12. SPH particles with their density below $n_{\rm crit} \approx 0.2$ cm^{-3} ionizes almost immediately as we find in the *r*-type phase. When an ionized particle is well above the cosmic mean density $n_{\rm mean} = 3.3 \times 10^{-4}$ cm^{-3}, the density of the particles drops close to $N_{\rm mean}$ within a few megayear. Particles with $n \gtrsim n_{\rm crit}$ are shielded regions and ionized later. Particles starting at a higher

density tend to ionize later since they are more deeply embedded in a minihalo.

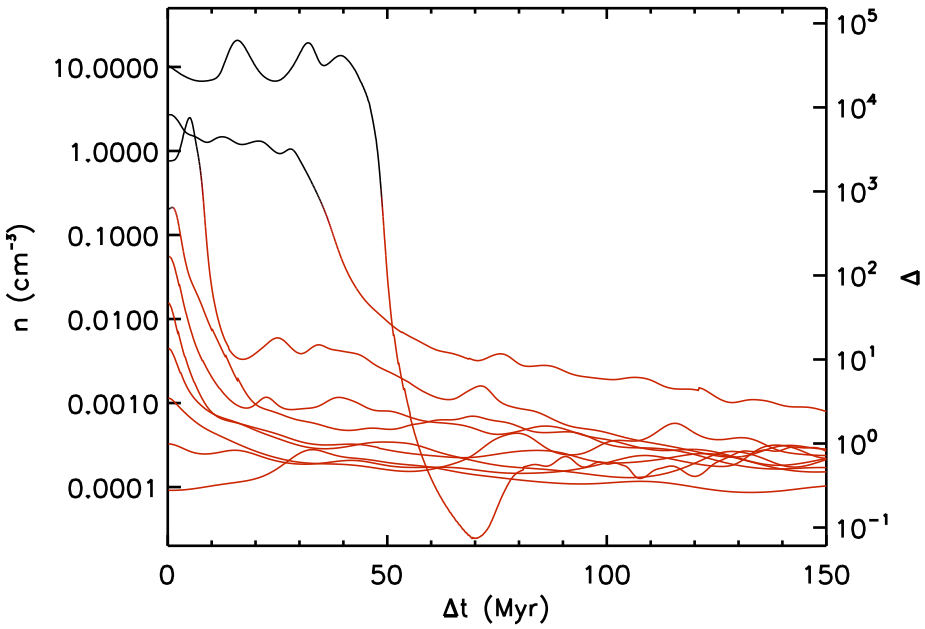

Figure 2.12: Densities of ten SPH particles in M_z10_I0 as functions of time. The color represents the ionization status of the particle with the black color meaning that neutral and red ionized.

2.3.2.1 Role of Shielding and Hydrodynamics

The dual phase evolution seen in this work is a consequence of taking into account both self-shielding of dense neutral gas and hydrodynamics. To bring out their importance, we run one case without shielding (M_I0_z10_NS) and another case without hydrodynamics (M_I0_z10_ND). The profiles of χ, P_M and dC_r/dn are shown for M_I0_z10_NS and M_I0_z10_ND in Figure 2.3, 2.4 and 2.10, respectively. The column density maps are provided in Fig-

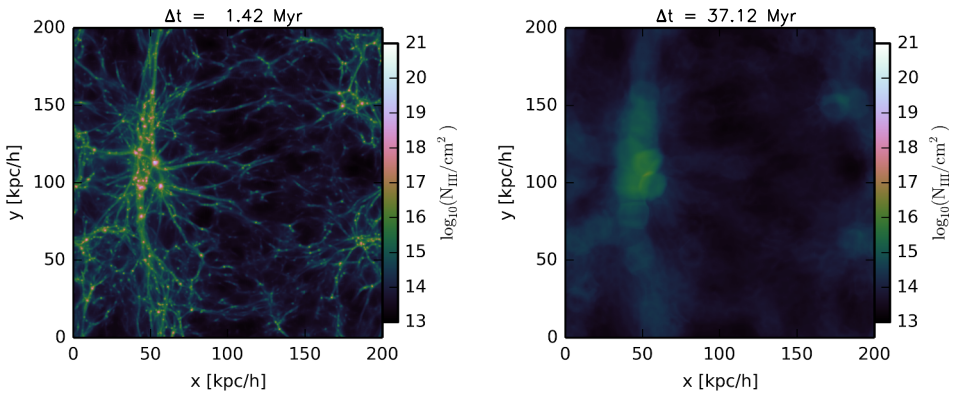

Figure 2.13: Projected neutral column density of M_I0_z10_NS at $\Delta t = 1.4$ Myr (left) and 37 Myr (right).

ure 2.13 (M_I0_z10_NS) and Figure 2.14 (M_I0_z10_ND) for a comparison with Figure 2.6 for M_I0_z10. The clumping factor as a function of Δt is shown as the dot-dashed line in the upper panel of Figure 2.5 for M_I0_z10_NS and as the dotted line in Figure 2.1 for M_I0_z10_ND.

M_I0_z10_NS shows a flash ionization of gas at all densities, thereby, completely skipping the r-type phase. Unlike in M_I0_z10, gas with $n \gtrsim 2$ cm^{-3} that is shielded in M_I0_z10 is both highly ionized and photo-heated in M_I0_z10_NS from early times as shown for $\Delta t = 1.4$ Myr. Therefore, there is no ionization front at all in this run. The column density map at $\Delta t = 0.14$ Myr and 1.4 Myr, clearly lacks most of the spurious high-column density regions with $N_{\rm HI} > 10^{19}$ cm^2 seen in M_I0_z10. At $\Delta t = 37$ Myr, M_I0_z10_NS does not have any gas with $n \gtrsim 0.3$ cm^{-3} that makes up neutral clumps in M_I0_z10 because all the gas has been subject to the hydrodynamical feedback

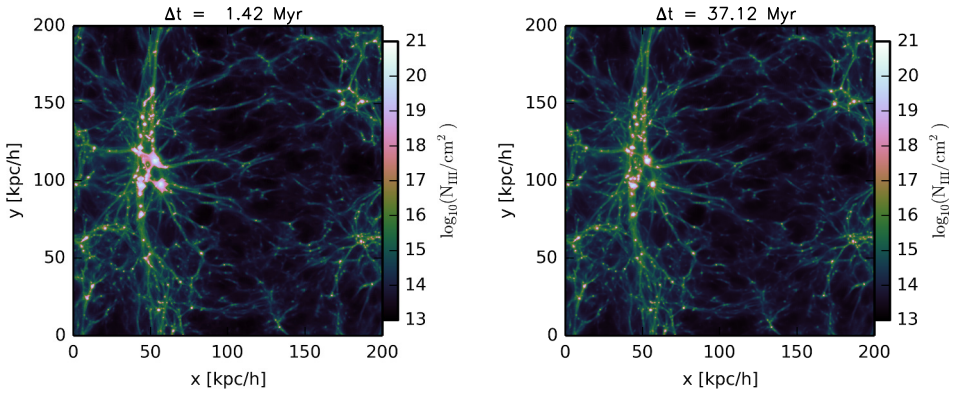

Figure 2.14: Projected neutral column density of M_I0_z10_ND at $\Delta t = 1.4$ Myr (left) and 37 Myr (right).

since EIBR was turned on. Thus, no HI clump appears in the column density map of M_I0_z10_NS at $\Delta t = 37$ Myr. The clumping factor in M_I0_z10_NS starts much higher than in M_I0_z10, but falls below that at $\Delta t \gtrsim 10$ Myr. At early times, dC_r/dn picks up a huge contribution from high density gas that should have been shielded as shown for $\Delta t = 1.4$ Myr and it lack contribution from gas evaporating from self-shielding clumps in the late time as shown for $\Delta t = 37$ Myr.

M_I0_z10_ND reproduces the r-type phase, but fails to turn into the d-type phase. Due to the absence of dynamics, the density pdf does not evolve except for a slight shift to lower density due to the Hubble expansion. High density ionized gas is not blown away thus maintaining high recombination rate. The simulation settles on an equilibrium state after a few megayear as ionization fronts can no longer propagate with the high recombination rate.

The column density map also shows lack of evolution between $\Delta t = 7.1$ Myr and $\Delta t = 37$ Myr although it shows almost identical evolution as in M_I0_z10 from $\Delta t = 0.14$ Myr to $\Delta t = 1.42$ Myr. The clumping factor grows in the same way as in M_I0_z10 up to $\Delta t \sim 1$ Myr. Then, the lack of evolution in P_M allows it grow further unlike in M_I0_z10 that it turns over due to the hydrodynamical feedback. After $\Delta t \gtrsim 3$ Myr, it converges a value as the simulation sets into an equilibrium.

2.3.3 Parameter Dependence of Clumping Factor

To study the dependence of the clumping factor on the intensity of EIBR (J_{21}) and on the redshift that the volume exposed to the EIBR (z_i), we run two cases, which the simulation parameters are same as in M_I0_z10 except that J_{21} is changed to 0.3 (M_I-0.5_z10) and 0.1 (M_I-1_z10) from 1 and two cases that z_i is changed to 9 (M_I0_z9) and 8 (M_I0_z8) from 10. In the upper panel of Figure 2.5, we compare the clumping factor for changing turn-on redshift of EIBR (z_i) and intensity of EIBR (J_{21}). While the r-type and the d-type phases show up globally, the overall magnitude of the clumping factor and the timing of the transition between the two phases differ.

For decreasing J_{21}, the clumping factor starts lower and turns over later. This is expected as ionization fronts with a lower intensity are expected to propagate more slowly and transition into the d-type phase at a lower density. This is clearly shown by $n_{\rm crit}$ in Figure 2.11. $n_{\rm crit}$ in lower intensity cases rises more slowly and settles at lower densities. In M_I-1_z10, $n_{\rm crit}$ settles

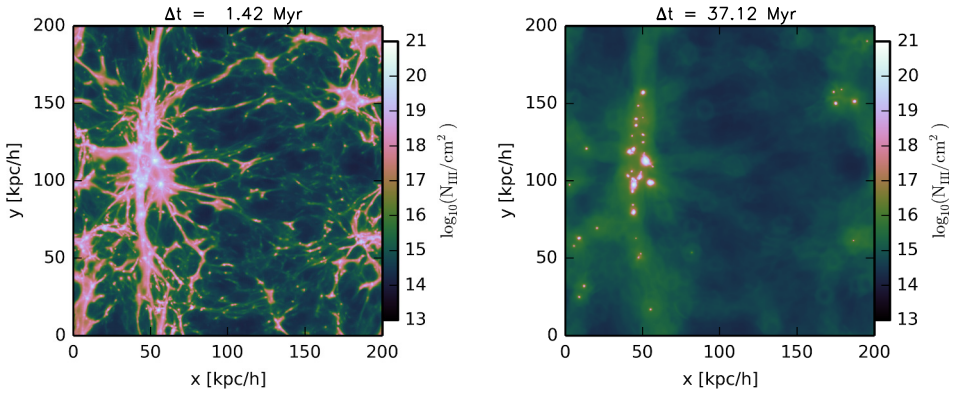

Figure 2.15: Projected neutral column density of M_I-1_z10 at $\Delta t = 1.4$ Myr (left) and 37 Myr (right).

at ~ 0.04 cm^{-3} while it does at ~ 0.2 cm^{-3} in M_I0_z10. Thus, gas with $0.04 \lesssim n \lesssim 0.2$ cm^{-3} does not contribute to the clumping factor in M_I-1_z10 as indicated by dC_r/dn in the bottom panel of Figure 2.8. The column density map of M_I-1_z10 (Fig. 2.15) at $\Delta t = 1.42$ Myr and 7.1 Myr shows that high column density regions with $N_{HI} \gtrsim 10^{19}$ cm^{-2} are more extended than in M_I0_z10 (Fig. 2.6) suggesting that the EIBR in M_I-1_z10 is not able to penetrate to as deeply as in M_I0_z10. The trend that lower intensities give lower clumping factors reverses at $\Delta t \sim 50$ Myr. This is because lower intensity cases retain more neutral clumps at late times that produce dense ionized gas from evaporation. The column density map of M_I-1_z10 at $\Delta t = 37$ Myr clearly shows more numerous neutral clumps then in M_I0_z10.

For a lower z_i, the clumping factor shows a similar time evolution, but the overall magnitude is significantly higher. We list the peak clumping

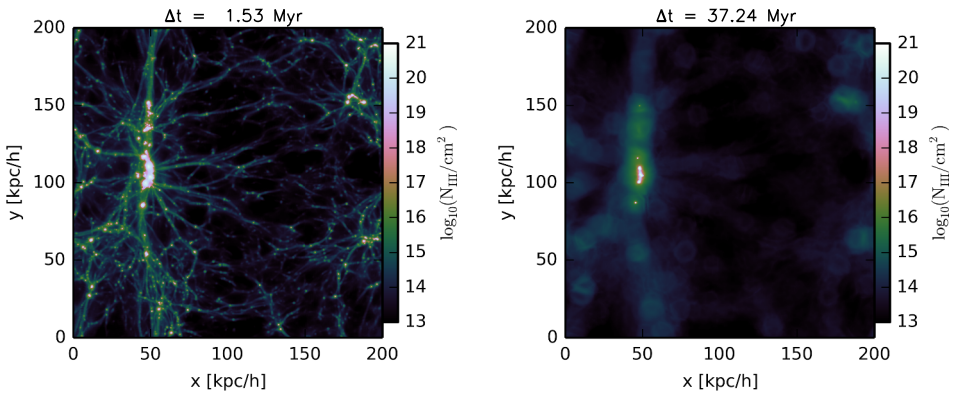

Figure 2.16: Projected neutral column density of M_I0_z8 at $\Delta t = 1.5$ Myr (left) and 37 Myr (right).

factor, $C_r^{\rm peak}$, in Table 2.1. The difference in the peak clumping factor mostly characterizes the difference at other times, too. The peak clumping factor almost scales as the inverse of the cosmic density giving $C_r^{\rm peak} \sim 21\left(\frac{1+z}{11}\right)^{-3}$. Note that the clumping factor is not only proportional to the recombination rate, but also to the inverse of the background rate from ionized gas with the cosmic mean density. Since recombination rate per mass goes as the cosmic mean density, this means the change of $C_r^{\rm peak}$ in lower z_i cases is mainly due to the Hubble expansion and not due to the actual change in the recombination rate.

To separate out the effect of cosmic expansion, we plot directly the recombination rate per hydrogen atom, $dN_{\rm rec}/dt$, in the lower panel of Figure 2.5. The $dN_{\rm rec}/dt$ at the peak is almost the same for different z_i's, but it falls more slowly for lower z_i's. Since lower z_i cases have more massive halos

at the high-mass end, they can have more dense ionized gas supplied by mini-halos in late time. Density pdf in the middle panel of Figure 2.9 shows that M_I0_z8 has more gas with $n > 0.1$ cm^{-3} than M_I0_z10 does. The column density map of M_I0_z8 (Fig. 2.16) shows that clumps are crowded in a smaller region than in M_I0_z10. This is also a sign that structures are more growth in M_I0_z8.

2.3.4 Comparison to Emberson et al. 2012 ([12])

Emberson et al. 2012 ([12]) report their result on the dependence of the clumping factor on the intensity and the timing of EIBR in the Figure 4 of their work[3]. Their clumping factor is larger than what recent works predict. For example, their clumping factor is well above 10 for $z_i < 10$ and $\Gamma_{-12} > 1$, which is expected to be a typical case during the EoR. Using parameters in Table 2.1, we can read expected clumping factor for each of our simulations. Since hydrodynamics not included in their methodology creates time dependence in the clumping factor result, our results are not directly comparable to their work expect for M_I0_z10_ND that we run without dynamics.

$z_i = 10$ and $\Gamma_{-12} = 9.2$ adopted in M_I0_z10_ND gives ~ 20 for the expected clumping factor in the left panel of Figure 4 of [12]. This is lower than 26 that C_r converges to in M_I0_z10_ND. We note, however, that their clumping factor corresponds to C_i in this work that neglects the effect of tem-

[3]Although both the left panel and the right panel give the clumping factor for given z_i and Γ_{-12}, each panel gives a somewhat different clumping factor for parameters of our interest. We shall adopt the left panel in this discussion.

perature enhancing the clumping factor. C_i converges to 21 in M_I0_z10_ND giving a good agreement with their result.

In the corresponding hydrodynamics-included run M_I0_z10, C_r peaks and turns over at a lower value, 21. For different z_i's, M_I0_z9 and M_I0_z8 give $C_r^{\rm peak}$=28 and 38 whereas their result predicts $\sim$ 30 and $\sim$ 45. And, for different Γ_{12}, M_I-0.5_z10 and M_I-1_z10 give $C_{r,\rm peak}$=13 and 8 whereas their result predicts $\sim$ 15 and $\sim$ 12. Given that C_i is even lower than C_r, our hydrodynamic simulations yields less clumpiness than in their work even when the clumpiness peaks. But, $C_{r,\rm peak}$ is still comparable to their result and is much higher than the results of other recent works giving less than ten.

2.3.5 Ionizing Photon Budget for Small-scale Structure

If one does not take into account the spatial variation of physical density and temperature in a volume like our simulation box, thereby assuming that density and temperature are homogeneous within the volume, the recombination rate per volume is given by $\alpha_B(\bar{T})\bar{n}_{\rm HII}\bar{n}_e$ ("background rate", hereafter), which underestimates the true rate by a factor of C_r as described in Equation 2.9. In the upper panel of Figure 2.17, we present the integrated recombination per H atom from the turn-on of EIBR, $N_{\rm rec}$, for M_I0_z10, M_I0_z9, M_I0_z8, M_I-0.5_z10 and M_I-1_z10. We also show the expected integrated recombination for the background rates in the same figure for comparison. Here, C_r gives the ratio of slope for each case and that of corresponding background case.

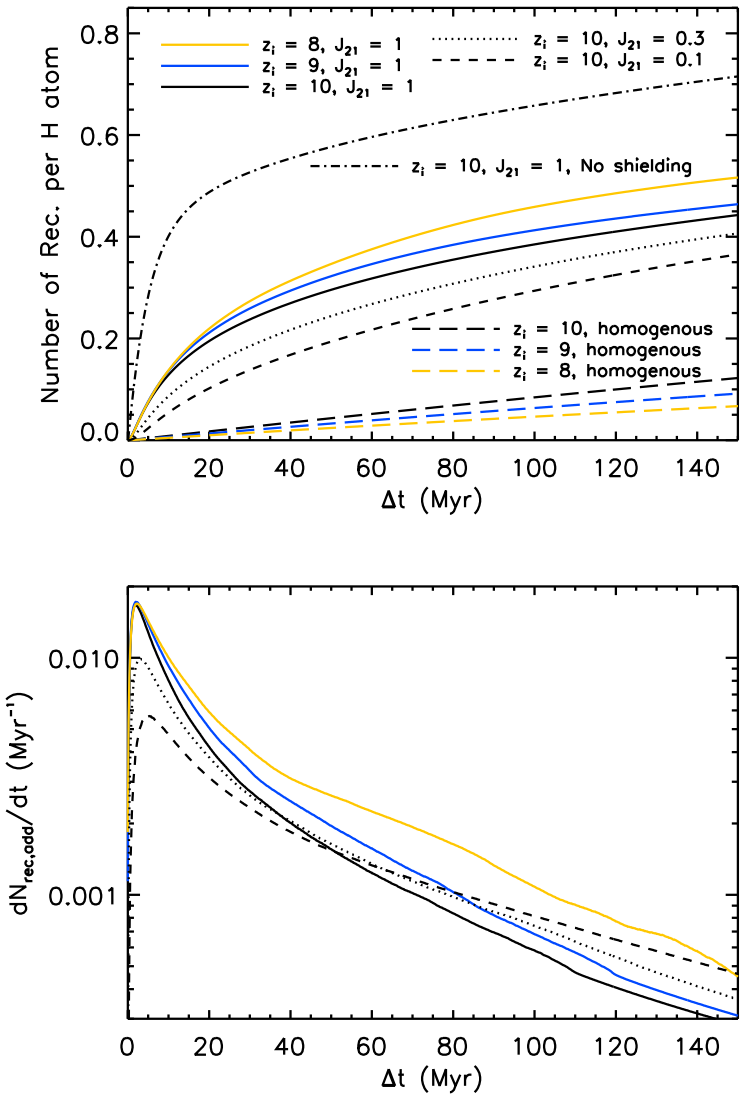

Figure 2.17: (upper) Accumulated recombination per H atom, $N_{\rm rec}$, for M_I0_z10 (black solid), M_I0_z9 (blue solid), M_I0_z8 (yellow solid), M_I-0.5_z10 (black dotted), M_I-1_z10 (black dashed) and M_I0_z10_NS (black dot-dashed). Black, blue and yellow long dashed line describing the "background" recombination rate calculated from the average gas density and the temperature for $z_i = 10,\ 9$ & 8, respectively. (lower) Instantaneous recombination rate minus the background rate, $dN_{\rm rec,add}/dt$. Line types represent the same cases as in the upper panel.

At early times ($\Delta t \lesssim 20$ Myr), high C_r makes the actual $N_{\rm rec}$ accumulate much faster than that of the background. As C_r decays toward unity later ($\Delta t \gtrsim 20$ Myr), the true $N_{\rm rec}$ goes in parallel to the background case. As a result, the difference in $N_{\rm rec}$ between the actual rate and the background rate becomes a fixed quantity at late times ($\Delta t \gtrsim 100$ Myr). We define the difference between the two cases in $N_{\rm rec}$ at $\Delta t = 150$ as $N_{\rm rec,add}$ and interpret it as *the additional ionizing photon budget for ionizing small-scales structures from the pre-ionization phase*. $N_{\rm rec,add}$ can be obtained by integrating $dN_{\rm rec,add}/dt = (C_r - 1)dN_{\rm rec}/dt$ from $t = 0$ to Δt. $N_{\rm rec,add}$ is listed in Table 2.1 for each simulation. We find that $N_{\rm rec,add}$ is larger for lower z_i's and higher J_{21}'s giving $\sim 0.3\ J_{21}^{0.15} \left(\frac{1+z}{11}\right)^{-1.6}$.

$dN_{\rm rec,add}/dt$ is shown in the lower panel of Figure 2.17. We fit the result with a 3rd order polynomial in log-log space as following.

$$\begin{aligned} \log\left(\frac{dN_{\rm rec,add}}{dt}\right) &= a_0 + a_1(\log\Delta t) + a_2(\log\Delta t)^2 \\ &+ a_3(\log\Delta t)^3 \qquad (\Delta t > 2 \text{ Myr}) \\ \\ &= 0 \qquad\qquad\qquad (\Delta t < 2 \text{ Myr}) \end{aligned} \tag{2.17}$$

For simplicity of fitting, we do not fit for $\Delta t < 2$Myr that the recombination rate is rising rapidly. Recombination during that time can be ignored with negligible error in the time scale of tens of Myr. The fitted parameters are provided in Table 2.2 for M_I0_z10, M_I0_z9, M_I0_z8, M_I-0.5_z10 and M_I-1_z10.

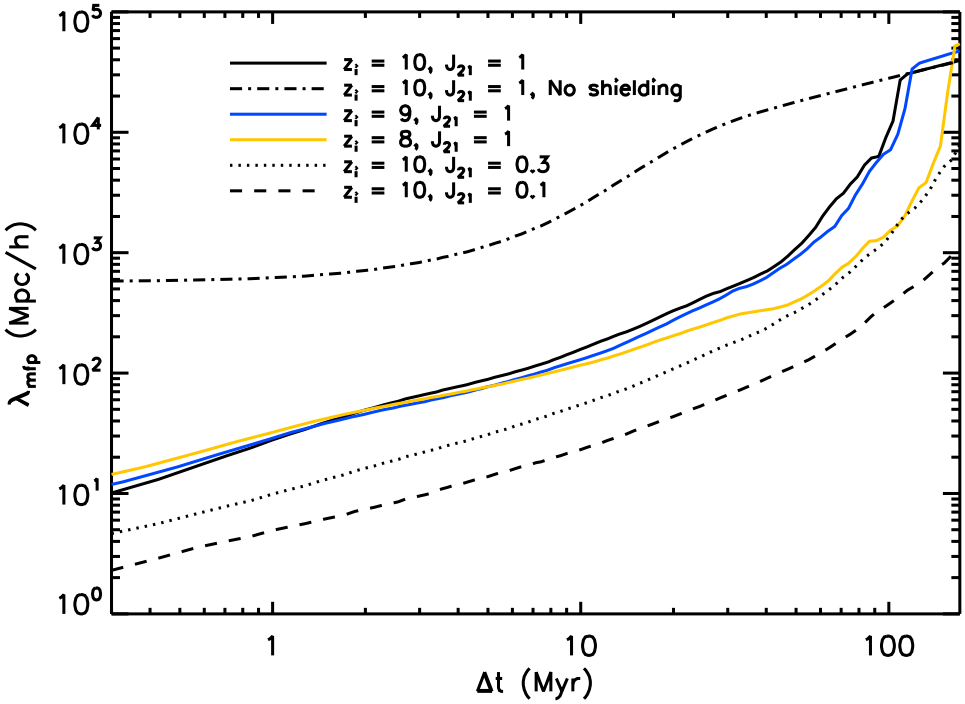

Figure 2.18: Mean free path plotted as a function of Δt for M_I0_z10 (black solid), M_I0_z9 (blue solid), M_I0_z8 (yellow solid), M_I-0.5_z10 (black dotted), M_I-1_z10 (black dashed) and M_I0_z10_NS (black dot-dashed).

2.3.6 Mean Free Path

The mean free path of UV photons ($\lambda_{\rm mfp}$, hereafter) can be used instead of recombination rate to model the effect of small-scale structures in large-scale EoR simulations. For example, [27] set $\lambda_{\rm mfp} = 80\ h^{-1}$ Mpc comoving in their simulation preventing UV photons from propagating further than $\lambda_{\rm mfp}$ from sources.

$\lambda_{\rm mfp}$ can be calculated using projected 2D HI column density maps like

those in Figure 2.6, Figure 2.13, Figure 2.14, Figure 2.15 and Figure 2.16. We first calculate the mean cross-section

$$\bar{\sigma} = \frac{\int_{13.6\mathrm{eV}/h_p}^{\infty} d\nu \frac{I_\nu}{h_p \nu} \sigma_\nu}{\int_{13.6\mathrm{eV}/h_p}^{\infty} d\nu \frac{I_\nu}{h_p \nu}} = 1.62 \times 10^{-18}\ \mathrm{cm}^2, \tag{2.18}$$

where h_p in the Planck's constant and I_ν is the intensity of EIBR. Thus, we can regard bright colored (pink and white) regions with $N_{\mathrm{HI}} \gtrsim 10^{18}\mathrm{cm}^{-2}$ as optically thick in the column density figures. The average $e^{-N_{\mathrm{HI}}\bar{\sigma}}$ in the column density map gives transmitivity of EIBR through the entire simulation volume, $F_{\mathrm{out}}/F_{\mathrm{in}}$. Then,

$$\lambda_{\mathrm{mfp}} = -\frac{l_{\mathrm{box}}}{\log\left(F_{\mathrm{out}}/F_{\mathrm{in}}\right)}, \tag{2.19}$$

where l_{box} is the box size of the simulation. We show λ_{mfp} for M_I0_z10, M_I0_z9, M_I0_z8, M_I-0.5_z10, M_I-1_z10 and M_I0_z10_NS in Figure 2.18. We provide polynomial fitting

$$\log(\lambda_{\mathrm{mfp}}) = b_0 + b_1 \log(\Delta t) + b_2 (\log \Delta t)^2 + b_3 (\log \Delta t)^3 \tag{2.20}$$

for our results that works up to $\lambda_{\mathrm{mfp}} = 2\ h^{-1}$ Gpc in Table 2.2. Note that it is meaningless to keep track of λ_{mfp} if it is well above a gigaparsec.

In general, λ_{mfp} has a strong time-dependence cautioning the idea of characterizing it with a single number. All cases other than M_I0_z10_NS shows two to three orders of magnitude change from $\Delta t = 1$ to 100 Myr. At $\Delta > 100\ h^{-1}$ Mpc, λ_{mfp} in those cases reach close to or exceed a gigaparsec that is larger than most EoR simulations run these days.

M_I0_z10_NS represents the case that optically thick clumps are taken out from M_I0_z10 and only tiny neutral fraction in diffuse ionized gas limits $\lambda_{\rm mfp}$. $\lambda_{\rm mfp}$ in M_I0_z10 converges to that in M_I0_z10_NS at $\Delta t \approx 100$ Myr when the last clump disappears from the simulation volume. $\lambda_{\rm mfp}$ in M_I0_z10_NS starts from 500 h^{-1} Mpc and grows above 10^4 h^{-1} Mpc as time goes. $\lambda_{\rm mfp}$ in M_I0_z10 stays roughly two orders of magnitude below that in M_I0_z10_NS while neutral clumps exist. This shows that neutral clumps are the major factor that limits $\lambda_{\rm mfp}$.

For lower z_i's with J_{21} fixed at 1, $\lambda_{\rm mfp}$ evolves similarly in the early time ($\Delta t < 10$ Myr), but rises more slowly in the late time ($\Delta t < 10$Myr). This is similar to the trend in recombination rate discussed above. Since the most massive halos in the simulation is more massive at lower z_i's, lower-z_i cases are more limited by neutral clumps in most massive halos in the late time.

For lower J_{21}'s with z_i fixed at 10, $\lambda_{\rm mfp}$ is lower at all Δt's as expected from slower evaporation of minihalos and higher equilibrium fraction of HI in diffuse IGM. Notable, the time dependence at $\Delta t < 10$ Myr stays nearly same with the global magnitude going as J_{21}. In the late time ($\Delta t < 10$ Myr), slower evaporation of neutral clumps in lower J_{21} cases seems to make $\lambda_{\rm mfp}$ rise more slowly. Due to the strong scaling of $\lambda_{\rm mfp}$ in J_{21}, $\lambda_{\rm mfp}$ stays below 100 h^{-1} Mpc for a much longer time (~ 50 Myr) in M_I-1_z10 then in M_I0_z10, M_I0_z9 or M_I0_z8 ($\lesssim 10$ Myr).

Table 2.2: Fitting Parameters for $\frac{dN_{\rm rec,add}}{dt}$ and $\lambda_{\rm mfp}$

label	z_i	J_{21} (Γ_{-12})	a_0	a_1	a_2	a_3	b_0	b_1	b_2	b_3
M_I0_z10	10	1 (9.2)	-3.8	-0.18	-0.097	-0.0078	3.4	0.70	-0.047	0.027
M_I-0.5_z10	10	0.3 (2.8)	-3.9	-0.67	0.14	-0.034	2.37	0.61	-0.023	0.025
M_I-1_z10	10	0.1 (0.92)	-5.2	0.094	-0.056	-0.012	1.6	0.60	-0.028	0.021
M_I0_z9	9	1 (9.2)	-3.5	-0.56	0.079	-0.030	3.38	0.59	-0.044	0.031
M_I0_z8	8	1 (9.2)	-4.0	-0.16	0.019	-0.027	3.5	0.57	-0.087	0.030

2.4 Numerical Convergence

How sensitive is our result to the size of the box? We adopt 200 h^{-1} kpc as a compromise between having as large size as possible and spending reasonable amount of computational time running multiple cases. To test the convergence of our result in the box-size, we run two cases, S_I0_z10 and L_I0_z10, that have the same z_i and J_{21} as in M_I0_z10, but have different box sizes. S_I0_z10 is run in a 100 h^{-1} kpc box and L_I0_z10 is run in a 400 h^{-1} kpc box. [12] claim 1 Mpc as the box-size of convergence, but their result in Figure 9 shows the clumping factor is underestimated only by 5% and 20% for 500 kpc and 250 kpc that are similar to the box-size of M_I0_z10 and L_I0_z10, respectively.

Clumping factors in S_I0_z10 (green solid), M_I0_z10 (black solid) and L_I0_z10 (red solid) are compared in Figure 2.1. The clumping factor in S_I0_z10 is smaller than in other cases by $\sim$ 20%. M_I0_z10 and L_I0_z10 agree with each other within a percent up to Δt = 2 Myr, but L_I0_z10 declines more slowly leading to a growing difference. At Δt = 15 Myr, the difference grows to $\sim$ 10%. The r-type phase seems to converge for M_I0_z10,

but the d-type phase seems to show some box-size effect albeit small. This can be understood from the fact that the clumping factor in the d-type phase is contributed from self-shielding neutral clumps hosted by minihalos. As time goes, neutral clumps with lower masses finish evaporating earlier and the clumping factor is contributed by evaporation from higher mass clumps that are likely to be hosted by minihalos with higher masses. Since more massive minihalos are more likely to be susceptible to the box-size effect, we expect to see growing difference at late time. In short, L_I0_z10 has more high-mass minihalos that can host larger neutral clumps and contribute to the clumping factor at late time.

2.5 Summary and Conclusion

We have explored recombination and mean free path of UV photons in a sub-Mpc volume after the volume is exposed to EIBR. By using baryonic particles with the mass $\approx 50 M_{\odot}$, we resolve minihalos and surrounding small-scale structures at mass scales down to 10^4 $M_{\odot}$. In order to take into account the effect of self-shielding dense clumps, we have coupled a shielding algorithm to hydrodynamics and tested the algorithm against an one-dimensional code for a spherically symmetric problem. For the sub-grid prescription for large-scale EoR simulations, we fit our result for recombination rate and mean free path with polynomials in Equation (2.17) and Equation (2.20), respectively, with the fitting parameters listed in Table 2.2.

As a result of resolving small-scale structures taking in account both

the hydrodynamics and radiative transfer, we find a dual phase evolution of the clumping factor characterized by the super-sonic r-type phase and the sub-sonic d-type phase. The clumping factor starts rising rapidly as super-sonic ionization fronts propagates toward higher densities. After a few megayears, the ionization fronts transition into the d-type as they become subsonic. The clumping factor peaks at the transition and begins to decay as the hydro-dynamical feedback expands ionized gas lowering its density. Dense clumps of gas self-shield themselves from EIBR gradually evaporating over time on the surfaces. Newly ionized gas from the clumps contributes to the clumping factor at late time.

Our clumping factor at the peak is comparable to or well above 10 as in [12]. Thus, we conclude that their high clumping factor is real although it is a transient phenomenon that lasts less than a few megayears. As the clumping factor converges to unity after tens of megayears, the recombination rate also converges to that of the background universe. But, the initial boost in the recombination due to high clumping factor makes a meaningful difference in the accumulated recombination that scales as $N_{\rm rec,add} \sim 0.3\ J_{21}^{0.15} \left(\frac{1+z_i}{11}\right)^{-1.6}$.

$\lambda_{\rm mfp}$ grows continuously cautioning the approach of using it as a fixed value in large-scale simulations. $\lambda_{\rm mfp}$ depends sensitively on the intensity of EIBR being nearly proportional to it. For cases with $J_{21} = 1$, $\lambda_{\rm mfp}$ grows to hundreds of Mpc in a few tens of Myr that is practically infinite in most EoR simulations. For $J_{21} = 0.1$, however, $\lambda_{\rm mfp}$ is below a hundred Mpc for ~ 30 Myr. Thus, it is important to know both when the volume is ionized and how

intense is the background when using $\lambda_{\rm mfp}$ in EoR simulations.

Chapter 3

The Kinetic Sunyaev-Zel'dovich Effect on the Temperature Anisotropy of Cosmic Microwave Background

The temperature of the CMB changes as free electrons in ionized gas Compton scatter CMB photons: the bulk peculiar velocity of electrons induces Doppler shifts in the energy of the CMB photons. While the spectrum of the CMB remains that of a black body, its temperature changes as following.

$$\frac{\Delta T(\hat{\gamma})}{T} = -\int d\tau e^{-\tau} \frac{\hat{\gamma} \cdot \mathbf{v}}{c}, \tag{3.1}$$

where $\hat{\gamma}$ is the line-of-sight unit vector, $\mathbf{v}$ the peculiar velocity field, and τ the optical depth to Thomson scattering integrated through the IGM from $z = 0$ to the surface of last scattering at $z_{\rm rec} \approx 10^3$, where

$$d\tau = c\ n_e(z)\sigma_{\rm T}(\frac{dt}{dz})dz. \tag{3.2}$$

Equation (3.1) can be rewritten in the following form:

$$\frac{\Delta T}{T}(\hat{\gamma}) = -\frac{\sigma_T \bar{n}_{e,0}}{c} \int \frac{ds}{a^2}\ e^{-\tau} \mathbf{q} \cdot \hat{\gamma}, \tag{3.3}$$

where $\mathbf{q} \equiv \chi \mathbf{v}(1+\delta)$ is the momentum of ionized gas, $\chi \equiv n_e/(n_{\rm H} + 2n_{\rm He})$ the ionization fraction, $\bar{n}_{e,0} = \bar{n}_{\rm H,0} + 2\bar{n}_{\rm He,0}$ the mean number density of electrons

at the (fully-ionized) present epoch, and s the distance photons travelled from a source to the observer in comoving units. As longitudinal momentum fields cancel out in the line-of-sight integral of Equation (3.3), the angular power spectrum of Equation (3.3), C_ℓ, at large multipoles is dominated by the power spectrum of the transverse momentum field, $P_{q_\perp}(k)$, and is given by [73, Also, see Chapter 3.1 for the derivation.]

$$C_\ell = \frac{1}{2}\left(\frac{\sigma_T \bar{n}_{e,0}}{c}\right)^2 \int \frac{ds}{s^2 a^4} e^{-2\tau} P_{q_\perp}\left(k = \frac{l}{s}, s\right). \tag{3.4}$$

3.1 Transverse Momentum Power Spectrum and kSZ Effect

3.1.1 Suppression of longitudinal modes

An important observation of the nature of kSZ is that it is given by the *transverse* (vector-mode or spin-1) momentum field, and the longitudinal contribution is suppressed. To show this, we Fourier transform Equation (3.3):

$$\frac{\Delta T}{T}(\hat{\gamma}) = -\frac{\sigma_T n_{e,0}}{c} \int \frac{ds}{a(s)^2} e^{-\tau} \int \frac{d^3k}{(2\pi)^3} [\hat{\gamma} \cdot \tilde{\mathbf{q}}(\mathbf{k}, s)] e^{-i\mathbf{k}\cdot(s\hat{\gamma})}. \tag{3.5}$$

Decomposing the momentum vector in Fourier space, $\tilde{\mathbf{q}}$, into the longitudinal component, $\tilde{q}_\parallel \equiv \tilde{\mathbf{q}} \cdot \hat{k}$, and the transverse component, $\tilde{q}_\perp \equiv |\tilde{\mathbf{q}} - \hat{k}(\tilde{\mathbf{q}} \cdot \hat{k})|$, we obtain

$$\begin{aligned} \frac{\Delta T}{T}(\hat{\gamma}) &= -\frac{\sigma_T n_{e,0}}{c} \int \frac{ds}{a(s)^2} e^{-\tau} \int \frac{d^3k}{(2\pi)^3} \\ &\quad \left[x\tilde{q}_\parallel(\mathbf{k}, s) + \cos(\phi_{\hat{q}} - \phi_{\hat{\gamma}})(1 - x^2)^{1/2} \tilde{q}_\perp(\mathbf{k}, s)\right] e^{-iksx}, \end{aligned} \tag{3.6}$$

where $x \equiv \hat{k} \cdot \hat{\gamma}$, and $\phi_{\hat{q}}$ and $\phi_{\hat{\gamma}}$ are the angles that $\mathbf{k}$ makes with $\tilde{\mathbf{q}}$ and $\hat{\gamma}$, respectively.

If the factor e^{iksx} oscillates much more rapidly than the other quantities, the integral over s will be small due to cancellation. Recalling that $a(s)$, $\tau(s)$, and $\tilde{\mathbf{q}}$ all vary over the Hubble length scale, kx should be much smaller than H/c in order to avoid the cancellation. Namely, either the wavelength should be longer than the Hubble length, i.e., $k \lesssim H/c$, or the mode should be nearly perpendicular to the line-of-sight direction, i.e., $x \approx 0$. The former does not contribute much because the amplitude of such a long-wavelength mode is small. Thus, only the modes that are perpendicular to the line-of-sight direction, $x \approx 0$, have a chance to contribute to the kSZ signal.

However, in this configuration, the longitudinal component of the momentum field is also perpendicular to the line-of-sight, and vanishes when taken a dot-product with the line-of-sight, i.e., $x\tilde{q}_{\parallel} \approx 0$. Therefore, only the transverse mode survives in the integral, giving

$$\frac{\Delta T}{T}(\hat{\gamma}) = -\frac{\sigma_T n_{e,0}}{c} \int \frac{ds}{a(s)^2} e^{-\tau} \int \frac{d^3k}{(2\pi)^3} \cos(\phi_{\hat{q}} - \phi_{\hat{\gamma}})(1-x^2)^{1/2} \tilde{q}_{\perp}(\mathbf{k}, s) e^{-iksx}. \tag{3.7}$$

3.1.2 Angular Power Spectrum

Here, we follow steps similar to those in Chapter 7.3 of [74] to derive the angular power spectrum of CMB fluctuations induced by the kSZ effect.

Spherical harmonic decomposition of Equation (3.7) gives

$$
\begin{aligned}
a_{lm} &= \int d^2\hat{\gamma}\; Y_l^{m*}(\hat{\gamma})\frac{\Delta T}{T}(\hat{\gamma}) \\
&= -\frac{\sigma_T n_{e,0}}{c}\int d^2\hat{\gamma}\; Y_l^{m*}(\hat{\gamma})\int \frac{ds}{a(s)^2}e^{-\tau}\int \frac{d^3k}{(2\pi)^3}\cos(\phi_{\hat{q}}-\phi_{\hat{\gamma}})(1-x^2)^{1/2}\tilde{q}_\perp(\mathbf{k},s)e^{-iksx} \\
&\equiv -\frac{\sigma_T n_{e,0}}{c}\int \frac{d^3k}{(2\pi)^3}\; f_{lm}(\mathbf{k}), \qquad (3.8)
\end{aligned}
$$

where

$$
\begin{aligned}
f_{lm}(\mathbf{k}) &\equiv \int d^2\hat{\gamma}\; Y_l^{m*}(\hat{\gamma})\int \frac{ds}{a(s)^2}e^{-\tau}\cos(\phi_{\hat{q}}-\phi_{\hat{\gamma}})(1-x^2)^{1/2}\tilde{q}_\perp(\mathbf{k},s)e^{-iksx} \\
&= \int d^2\hat{\gamma}\; Y_l^{m*}(\hat{\gamma})\int \frac{ds}{a(s)^2}e^{-\tau}\cos(\phi_{\hat{q}}-\phi_{\hat{\gamma}})(1-x^2)^{1/2}\tilde{q}_\perp(\mathbf{k},s) \\
&\quad \times 4\pi\sum_{LM}(-i)^L j_L(ks)Y_L^M(\hat{\gamma})Y_L^{M*}(\hat{k}). \qquad (3.9)
\end{aligned}
$$

We first choose a convenient coordinate system in which the z-direction lies on that of the mode vector, i.e., $\hat{k}=\hat{z}$, and the azimuthal direction is the same as the direction of the momentum vector, i.e., $\phi_{\hat{q}}=0$. In this case,

$Y_L^{M*}(\hat{k})$ simplifies to $Y_L^{M*}(\hat{z}) = \delta_{M0}\sqrt{\frac{2L+1}{4\pi}}$, giving

$$f_{lm}(k\hat{z}) = \sqrt{4\pi}\int \frac{ds}{a(s)^2} e^{-\tau}\tilde{q}_\perp(\mathbf{k},s)\sum_L (-i)^L\sqrt{2L+1}j_L(ks)$$

$$\int d^2\hat{\gamma}\; Y_L^0(\hat{\gamma})\;\cos\phi\;\sin\theta\; Y_l^{m*}(\hat{\gamma})$$

$$= \sqrt{\frac{8\pi^2}{3}}\int \frac{ds}{a(s)^2} e^{-\tau}\tilde{q}_\perp(\mathbf{k},s)\sum_L (-i)^L\sqrt{2L+1}j_L(ks)$$

$$\int d^2\hat{\gamma}\; Y_L^0(\hat{\gamma})\;\left[Y_1^{-1}(\hat{\gamma}) - Y_1^1(\hat{\gamma})\right] Y_l^{m*}(\hat{\gamma}), \tag{3.10}$$

where θ and $\phi = \phi_{\hat{\gamma}}$ determine the line-of-sight vector as $\hat{\gamma} = (\cos\theta\sin\phi, \sin\theta\sin\phi, \cos\phi)$.

The integral over $\hat{\gamma}$ can be computed using,

$$\int d^2\hat{\gamma}\; Y_L^M(\hat{\gamma})\; Y_\Lambda^\mu(\hat{\gamma})\; Y_l^{m*}(\hat{\gamma}) = \sqrt{\frac{(2\Lambda+1)(2l+1)}{4\pi(2L+1)}}$$

$$\times C_{l\Lambda}(L,M;m,-\mu)C_{l\Lambda}(L,0;0,0)\delta_{M,m-\mu}, \tag{3.11}$$

where $C_{l\Lambda}(L,M;m,\mu)$ is the Clebsch-Gordan coefficient for adding the angular momentum quantum numbers (l,m) and (Λ,μ) and for forming (L,M). In our case, we have

$$f_{l,m=\pm1}(k\hat{z}) = \sqrt{2\pi(2l+1)}\int \frac{ds}{a(s)^2} e^{-\tau}\tilde{q}_\perp(\mathbf{k},s)\sum_L (-i)^L j_L(ks)\times$$

$$[\;\pm C_{l1}(L,0;\pm1,\mp1)C_{l1}(L,0;0,0))\;]\,. \tag{3.12}$$

Thus, the relevant coefficients are

$$C_{l1}(l+1,0;0,0) = \sqrt{\frac{l+1}{2l+1}},$$

$$C_{l1}(l+1,0;\pm 1,\mp 1) = \sqrt{\frac{l}{2(2l+1)}},$$

$$C_{l1}(l,0;0,0) = 0,$$

$$C_{l1}(l-1,0;0,0) = \sqrt{\frac{l}{2l+1}},$$

$$C_{l1}(l-1,0;\pm 1,\mp 1) = \sqrt{\frac{l+1}{2(2l+1)}}. \tag{3.13}$$

Putting these together gives

$$\begin{aligned} f_{l,m=\pm 1}(k\hat{z}) &= (-i)^{l+1}\sqrt{\frac{\pi l(l+1)}{2l+1}}\int \frac{ds}{a(s)^2} e^{-\tau}\tilde{q}_{\perp}(\mathbf{k},s)\left[j_{l+1}(ks)+j_{l-1}(ks)\right] \\ &= (-i)^{l+1}\sqrt{\pi l(l+1)(2l+1)}\int \frac{ds}{a(s)^2} e^{-\tau}\tilde{q}_{\perp}(\mathbf{k},s)\frac{j_l(ks)}{ks}. \end{aligned} \tag{3.14}$$

Now, we get back to the observer's frame by applying the standard rotation operator, $S(\hat{q})$, that takes the z-direction into $\hat{k}$. This gives

$$f_{lm}(\mathbf{k}) = \sum_{m'=\pm 1} D^l_{m,m'}(S(\hat{k}))f_{lm'}(k\hat{z}), \tag{3.15}$$

where $D^l_{mm'} = \langle l,m'|S|l,m\rangle$ is the matrix representation of the finite rotation

of an initial state (l, m) into a final state (l, m'). We obtain

$$a_{lm} = -\frac{\sigma_T n_{e,0}}{c} \int \frac{d^3k}{(2\pi)^3} \sum_{m'=\pm 1} D^l_{m,m'}(S(\hat{k})) \\ (-i)^{l+1}\sqrt{\pi l(l+1)(2l+1)} \int \frac{ds}{a(s)^2} e^{-\tau} \tilde{q}_\perp(\mathbf{k}, s) \frac{j_l(ks)}{ks}. \quad (3.16)$$

Finally, we calculate the angular power spectrum from $\langle a_{lm} a^*_{l'm'} \rangle = C_l \delta_{ll'} \delta_{mm'}$ and obtain

$$C_l = \frac{l(l+1)}{\pi} \left(\frac{\sigma_T n_{e,0}}{c}\right)^2 \int \frac{ds}{a(s)^2} e^{-\tau(s)} \int \frac{ds'}{a(s')^2} e^{-\tau(s')} \int k^2 dk \frac{j_l(ks)}{ks} \frac{j_l(ks')}{ks'} P_{q_\perp}(k, s), \quad (3.17)$$

where $P_{q_\perp}$ is the power spectrum of $\tilde{\mathbf{q}}_\perp$ defined by $(2\pi)^3 P_{q_\perp}(k) \delta^D(\mathbf{k} - \mathbf{k}') = \langle \tilde{\mathbf{q}}_\perp(\mathbf{k}) \tilde{\mathbf{q}}^*_\perp(\mathbf{k}') \rangle$. Here, we have used the identity,

$$\int d^2\hat{k}\ D^l_{m,\pm 1}(S(\hat{k})) D^{l'*}_{m',\pm 1}(S(\hat{k})) = \frac{4\pi}{2l+1} \delta_{mm'} \delta_{ll'}. \quad (3.18)$$

The integral over k can be performed with Limber's approximation: when a function $g(k, s)$ varies much more slowly than the spherical Bessel function, one can approximate the integral as

$$\int k^2 dk j_l(ks) j_l(ks') g(k, s) \approx \frac{\pi}{2} \frac{\delta^D(s - s')}{s^2} g\left(k = \frac{l}{s}, s\right). \quad (3.19)$$

With this approximation, we finally obtain the desired formula for the kSZ power spectrum:

$$C_l = \left(\frac{\sigma_T n_{e,0}}{c}\right)^2 \int \frac{ds}{s^2 a(s)^4} e^{-2\tau(s)} \frac{P_{q_\perp}(k = l/s, s)}{2}. \quad (3.20)$$

This is Equation (3.4).

3.2 The Impact of Nonlinear Structure Formation on the Power Spectrum of Transverse Momentum Fluctuations and the Kinetic Sunyaev-Zel'dovich Effect

[1]

For the post-reionization signal, IGM is fully ionized and χ does not fluctuate to a good approximation. We thus need to model the density and velocity fluctuations of gas. When $\chi = 1$, the momentum field is given by $\mathbf{q} = \mathbf{v}(1+\delta)$. Since the velocity field is purely longitudinal in the linear regime, the transverse momentum field, $\mathbf{q}_\perp$, is given by $\mathbf{q}_\perp = (\mathbf{v}\delta)_\perp$ at leading order. The power spectrum is then given by the four-point function of two densities and velocities. Schematically, it is given by $\langle qq \rangle = \langle vv \rangle \langle \delta\delta \rangle + 2\langle v\delta \rangle^2 + \langle v\delta v\delta \rangle_c$. The last term is called the connected four-point term, while the others are the unconnected ones.

In the previous work, the connected term has been ignored. For example, the earlier analytical studies ignore non-linearity in density or velocity, and calculate only the unconnected terms using linear perturbation theory [48, 73, 11, 35]. An analytical model of [26] is still based upon linear theory for velocity, and ignores the connected term, but replaces the linear density power spectrum with a model for the non-linear density power spectrum by [53]. [40] use a similar approach with a halo model for the non-linear density power spectrum and argue that the connected term is negligible at large k. While we broadly agree with this conclusion, my aim is to quantify the

[1] This chapter is based on a published work in [50].

contribution of the connected term at large k, and also clarify the role of the connected term in obtaining the correct small-k limit of the transverse momentum power spectrum in perturbation theory.

Some of the previous "numerical" calculations of the post-reionization kSZ power spectrum [79, 64] still rely on the above analytical model that ignores the connected term, but takes the ingredient of the model, i.e., non-linear gas density power spectrum, from hydrodynamical simulations. Therefore, quantifying better the contribution of the connected term affects the results from the previous numerical work as well. [69] computed the kSZ power spectrum directly from their simulation and did not rely on the model.

Throughout this dissertation, we shall assume that gas traces dark matter. This is not a great approximation: shocks in the IGM generated by structure formation heat gas to high temperatures [e.g., 9], and the resulting gas pressure makes gas less clustered than dark matter particles. This effect on the kSZ power spectrum is modest at $\ell = 3000$ [64, 26]. Star formation converts gas into stars, further reducing the kSZ effect. Shaw et al. (2012) find that $D^{\rm kSZ}_{\ell=3000}$ can be suppressed by as much as 33% of the simulation without gas cooling and star formation. My goal in this dissertation is to quantify the error we make by ignoring the connected four-point term in the transverse momentum power spectrum. While my dark-matter-only results cannot be extrapolated to gas, we expect that a similar conclusion would still apply to gas, at least qualitatively.

3.2.1 Simulation

Throughout Section 3.2, we shall use a cosmological N-body simulation of collisionless particles using the "CubeP^{3}M" N-body code [24]. The simulation is run with 3456^3 particles in a comoving box of 1 h^{-1} Gpc on a side and is started at $z = 150$ using the initial condition generated using the Zel'dovich approximation and initial density power spectrum from the publicly available code CAMB [39]. This simulation was previously presented in [**?**].

The resolution of this simulation allows us to sample on average ~ 40 particles per $(\mathrm{Mpc}/h)^3$. This resolution allows us to avoid sampling artifacts in the velocity power spectrum up to $k \sim 1\ h\ \mathrm{Mpc}^{-1}$ [see Figure 3 of 80]. We then adaptively smooth particles to a grid of $(1008)^3$ cells [2]. Therefore, our simulation covers a dynamic range of $0.006 \lesssim k \lesssim 1\ h\ \mathrm{Mpc}^{-1}$ in wavenumber. The background cosmology is based on the *WMAP* 5-year data combined with constraints from baryonic acoustic oscillations, from observations of galaxies and large-scale structures, and from high-redshift Type Ia supernovae [$\Omega_M = 0.279, \Omega_\Lambda = 0.721, h = 0.701, \Omega_b = 0.0462, \sigma_8 = 0.8, n_s = 0.96$; 38].

[2] This adaptive smoothing is by an approach similar to that used in Smoothed Particle Hydrodynamics. In this case, spherical smoothing kernels are assigned to each particle, with their smoothing lengths adjusted so as to enclose the locations of the 32 nearest-neighbor particles. The mass per particle assigned to a given grid cell then corresponds to the fraction of its finite kernel volume which overlaps the cell volume.

3.2.2 Basics

In the post-reionization era, helium atoms are singly ionized until He II reionization occurs. We assume that hydrogen reionization finished at $z = 6$ and He II reionization occurred instantaneously at $z = 3$; thus, $\chi = 0.93$ for $3 < z < 6$ and $\chi = 1$ for $z \leq 3$. For the rest of the paper, we shall drop χ for notational simplicity and write $\mathbf{q} = \mathbf{v}(1+\delta)$. Therefore, the momentum power spectrum below should be rescaled by χ^2 to yield the correct value needed for computing the kSZ power spectrum in Chapter 3.2.6.

We start by Fourier transforming the momentum field, $\mathbf{q} = \chi\mathbf{v}(1+\delta)$, as

$$\tilde{\mathbf{q}}(\mathbf{k}) = \tilde{\mathbf{v}}(\mathbf{k}) + \int \frac{d^3k'}{(2\pi)^3}\tilde{\delta}(\mathbf{k}-\mathbf{k}')\tilde{\mathbf{v}}(\mathbf{k}'), \tag{3.21}$$

where $\tilde{\mathbf{q}}(\mathbf{k}) \equiv \int d^3\mathbf{x}\; e^{i\mathbf{k}\cdot\mathbf{x}}\mathbf{q}(\mathbf{x})$, etc. Then, its power spectrum is defined by

$$\begin{aligned}
& (2\pi)^3 P_{qq}^{ij}(\mathbf{k_1})\delta_D(\mathbf{k_1}+\mathbf{k_2}) \equiv \left\langle \tilde{q}^i(\mathbf{k_1})\tilde{q}^j(\mathbf{k_2})\right\rangle \\
= \;& \left\langle \tilde{v}^i(\mathbf{k_1})\tilde{v}^j(\mathbf{k_2})\right\rangle + 2\int \frac{d^3k'}{(2\pi)^3}\left\langle \tilde{\delta}(\mathbf{k_1}-\mathbf{k}')\tilde{v}^i(\mathbf{k}')\tilde{v}^j(\mathbf{k_2})\right\rangle \\
+ \;& \int \frac{d^3k'}{(2\pi)^3}\int \frac{d^3k''}{(2\pi)^3}\left\langle \tilde{\delta}(\mathbf{k_1}-\mathbf{k}')\tilde{v}^i(\mathbf{k}')\tilde{\delta}(\mathbf{k_2}-\mathbf{k}'')\tilde{v}^j(\mathbf{k}'')\right\rangle,
\end{aligned} \tag{3.22}$$

where the indices, i and j, denote the i'th and j'th components of the vector, respectively. The power spectrum of the transverse mode, $\tilde{\mathbf{q}}_\perp \equiv \tilde{\mathbf{q}} - \hat{\mathbf{k}}[\tilde{\mathbf{q}}\cdot\hat{\mathbf{k}}]$, is given by $P_{q_\perp} \equiv \Sigma_i P_{qq}^{ii}[1-(\hat{k}^i)^2]$. In linear and quasi-linear regimes ($k \lesssim 1\ h/\mathrm{Mpc}$), the velocity field is longitudinal (i.e. $\tilde{\mathbf{v}}(\mathbf{k}) = \hat{\mathbf{k}}\tilde{v}$) to a good

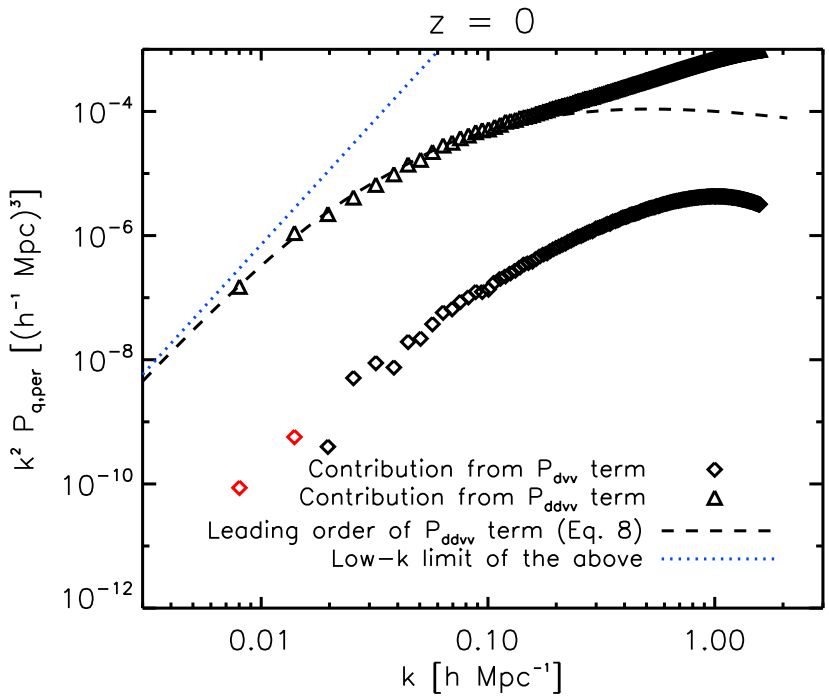

Figure 3.1: Contributions to $k^2 P_{q_\perp}$ from $\langle \delta vv \rangle$ (diamonds) and $\langle \delta\delta vv \rangle$ (triangles) in Equation (3.22) measured from the simulation at $z = 0$. The red triangles show negative values. The dashed line shows the lowest order calculation for the $\langle \delta\delta vv \rangle$ term, while the dotted line shows its low-k limit given in Equation (3.28).

approximation, which implies that $\langle vv \rangle$ and $\langle \delta vv \rangle$ in Equation (3.22) vanish. Figure 3.1 shows that the contribution to $P_{q_\perp}$ from $\langle \delta vv \rangle$ (diamonds) is a few orders of magnitude smaller than $\langle \delta v \delta v \rangle$ (triangles). The contribution from $\langle vv \rangle$ (not shown) is even smaller than $\langle \delta vv \rangle$ by another two orders of magnitude.

Keeping only the last term in Equation (3.22), we obtain

$$P_{q_\perp}(k) = \int \frac{d^3k'}{(2\pi)^3} \int \frac{d^3k''}{(2\pi)^3} [\hat{\mathbf{k}'} \cdot \hat{\mathbf{k}''} - (\hat{\mathbf{k}} \cdot \hat{\mathbf{k}'})(\hat{\mathbf{k}} \cdot \hat{\mathbf{k}''})] P_{\delta\delta vv}(\mathbf{k} - \mathbf{k}', -\mathbf{k} - \mathbf{k}'', \mathbf{k}', \mathbf{k}''), \tag{3.23}$$

where $P_{\delta\delta vv}$ is defined by

$$\begin{aligned} & (2\pi)^3 P_{\delta\delta vv}(\mathbf{q_1}, \mathbf{q_2}, \mathbf{q_3}, \mathbf{q_4}) \delta_D(\mathbf{q_1} + \mathbf{q_2} + \mathbf{q_3} + \mathbf{q_4}) \\ \equiv & \left\langle \tilde{\delta}(\mathbf{q_1}) \tilde{\delta}(\mathbf{q_2}) \tilde{v}^i(\mathbf{q_3}) \tilde{v}^j(\mathbf{q_4}) \right\rangle . \end{aligned} \tag{3.24}$$

3.2.3 Linear Regime

In the linear regime, we consider only the first order terms of $\tilde{\delta}$'s and $\tilde{v}$'s in $\langle \delta\delta vv \rangle$. Then, Gaussianity of linear v and δ fields yields

$$\begin{aligned} & \left\langle \tilde{\delta}(\mathbf{k_1} - \mathbf{k}') \tilde{v}^i(\mathbf{k}') \tilde{\delta}(\mathbf{k_2} - \mathbf{k}'') \tilde{v}^j(\mathbf{k}'') \right\rangle \\ = & \left\langle \tilde{\delta}(\mathbf{k_1} - \mathbf{k}') \tilde{v}^i(\mathbf{k}') \right\rangle \left\langle \tilde{\delta}(\mathbf{k_2} - \mathbf{k}'') \tilde{v}^j(\mathbf{k}'') \right\rangle \\ + & \left\langle \tilde{\delta}(\mathbf{k_1} - \mathbf{k}') \tilde{\delta}(\mathbf{k_2} - \mathbf{k}'') \right\rangle \left\langle \tilde{v}^i(\mathbf{k}') \tilde{v}^j(\mathbf{k}'') \right\rangle \\ + & \left\langle \tilde{\delta}(\mathbf{k_1} - \mathbf{k}') \tilde{v}^j(\mathbf{k}'') \right\rangle \left\langle \tilde{\delta}(\mathbf{k_2} - \mathbf{k}'') \tilde{v}^i(\mathbf{k}') \right\rangle . \end{aligned} \tag{3.25}$$

The first term in the above vanishes and the other terms lead to [40]

$$P_{q_\perp}(k, z) = \int \frac{d^3k'}{(2\pi)^3} (1 - \mu'^2) \left[\frac{1}{k'^2} P_{\delta\delta}(|\mathbf{k} - \mathbf{k}'|) P_{\theta\theta}(k') - \frac{1}{|\mathbf{k} - \mathbf{k}'|^2} P_{\delta\theta}(|\mathbf{k} - \mathbf{k}'|) P_{\delta\theta}(k') \right], \tag{3.26}$$

where $\theta \equiv \nabla \cdot \mathbf{v}$.

The linear velocity is related to the linear density by $\dot{a} f \tilde{\theta} = \tilde{\delta}$, where $f \equiv d \ln \delta / d \ln a$ and $a(t)$ is the Robertson-Walker scale factor. This gives the lowest order (LO) expression as [73]

$$P_{q_\perp}^{\rm LO}(k,z) = \dot{a}^2 f^2 \int \frac{d^3k'}{(2\pi)^3} P_{\delta\delta}^{lin}(|\mathbf{k}-\mathbf{k}'|,z) P_{\delta\delta}^{lin}(k',z) \times \frac{k(k-2k'\mu')(1-\mu'^2)}{k'^2(k^2+k'^2-2kk'\mu')}, \quad (3.27)$$

where $P_{\delta\delta}^{lin}$ is the linear matter density power spectrum. A similar derivation for the power spectrum of the sum of longitudinal and transverse momentum fields is presented in [49].

Taking $k \to 0$ limit of Equation (3.27), we find

$$P_{q_\perp}^{\rm LO}(k,z) \to \frac{8}{15} \dot{a}^2 f^2 k^2 \int \frac{dk'}{(2\pi)^2} \frac{[P_{\delta\delta}^{lin}(k',z)]^2}{k'^2}. \quad (3.28)$$

The k dependence of $P_{q_\perp}^{\rm LO}$ is thus given simply by k^2, which is independent of cosmology or the initial power spectrum. We shall discuss the physical origin of this dependence in Chapter 3.2.5.2.

We use the publicly available code CAMB [39] to compute $P_{\delta\delta}^{lin}$. The dashed line in Figure 3.1 shows $P_{q_\perp}^{\rm LO}$ at $z = 0$, while the dotted line shows the low-k limit (Eq. 3.28; Derived in Appendix 2.1).

3.2.4 Non-linear Regime

As Figure 3.1 shows, non-linear contributions to $P_{q_\perp}$ become dominant over the LO contribution at $k \gtrsim 0.2$ h/Mpc at $z = 0$, in agreement with the previous work [26, 40]. The current popular model of the post-reionization

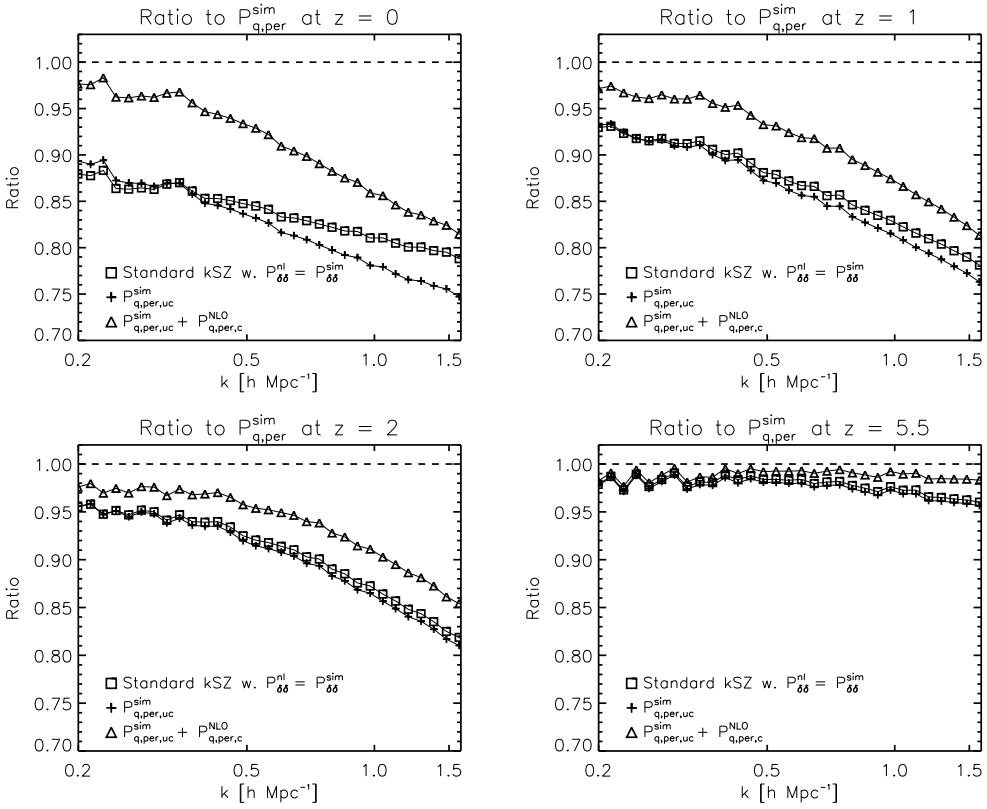

Figure 3.2: Ratios of various models for $P_{q_\perp}$ to $P_{q_\perp}$ directly measured from the simulation at $z = 0$ (upper left), $z = 1$ (upper right), $z = 2$ (lower left) and $z = 5.5$ (lower right). The models include Equation (3.29) with $P^{nl}_{\delta\delta}$ from the simulation (squares), Equation (3.26) with $P_{\delta\delta}$, $P_{\delta\theta}$ and $P_{\theta\theta}$ from the simulation (crosses), and the connected term calculated from perturbation theory added to the crosses (triangles).

kSZ signal uses an approximate expression for the transverse momentum power spectrum due to [26], which replaces one of $P_{\delta\delta}^{lin}$ in Equation (3.27) with the non-linear power spectrum, $P_{\delta\delta}^{nl}$:

$$P_{q_\perp}^{\rm S}(k,z) = \dot{a}^2 f^2 \int \frac{d^3k'}{(2\pi)^3} P_{\delta\delta}^{nl}(|\mathbf{k}-\mathbf{k}'|,z) P_{\delta\delta}^{lin}(k',z) \times \frac{k(k-2k'\mu')(1-\mu'^2)}{k'^2(k^2+k'^2-2kk'\mu')}. \quad (3.29)$$

This model modifies the second term in the square bracket of Equation (3.26) in the following way:

$$P_{\delta\theta}(|\mathbf{k}-\mathbf{k}'|)P_{\delta\theta}(k') = P_{\delta\delta}(|\mathbf{k}-\mathbf{k}'|)P_{\theta\theta}(k'). \quad (3.30)$$

This holds in the linear regime, and the entire term vanishes in the large k regime due to the pre-factor, $|\mathbf{k}-\mathbf{k}'|^{-2}$. In addition, the model approximates the velocity power spectrum by linear theory, i.e., $P_{\theta\theta} = P_{\delta\delta}^{lin}$. In this way, the model avoids having to model the velocity power spectrum and the density-velocity cross power spectrum, which is relatively poorly understood. We shall refer to this model as the *Standard* model (hence the superscript "S" in Eq. 3.29) in this paper.

However, Equation (3.26), which the Standard model aims to model, is not the full expression for $P_{q_\perp}$ because it neglects the contribution from the connected term, $\langle \delta v \delta v \rangle_c$. In the non-linear regime, non-linear growth makes both δ and v non-Gaussian, and thus there is no reason to think that the connected four-point term is negligible. We shall quantify the importance of this term in detail in Chapter 3.2.5.2.

3.2.5 Revisiting Non-linear Transverse Momentum Power Spectrum Model

3.2.5.1 Unconnected Term

In this chapter, we revisit accuracy of Equation (3.29). As noted above, this model is an approximation for the unconnected term (Eq. 3.26). Therefore, it, by design, does not take into account the connected term.

We test whether Equation (3.29) successfully approximates Equation (3.26) by evaluating it using the density power spectrum from the simulation, i.e., $P_{\delta\delta}^{nl} = P_{\delta\delta}^{sim}$, and compare it with Equation (3.26) using $P_{\delta\delta}$, $P_{\delta\theta}$, and $P_{\theta\theta}$ from the simulation.

In Figure 3.2, we show the ratios of Equation (3.29) (squares) and Equation (3.26) (crosses) to $P_{q_\perp}$ measured directly from the simulation. At $0.1 \lesssim k \lesssim 0.5$ h Mpc^{-1} at $z = 0$ (top left panel), the Standard model reproduces the unconnected term with high accuracy ($\sim 2\%$ level). However, the Standard model overestimates the unconnected term at larger k, reaching $\sim 5\%$ level at $k \sim 1$ h Mpc^{-1}. We attribute this error to the linear velocity assumption overestimating magnitudes of velocity modes. In the high-k limit, Equation (3.26) converges to [26]

$$P_{q_\perp}(k) = \frac{2}{3} P_{\delta\delta}(k) v_{rms}^2(k), \tag{3.31}$$

where $v_{rms}^2(k) = \int_{k' \leq k} \frac{d^3k'}{(2\pi)^3} P_{vv}(k')$ is the velocity dispersion. Non-linear correction makes the velocity power spectrum in the relevant k range *smaller* than the linear velocity power spectrum [see, e.g., Figure 1 of 54]; thus, linear

theory overestimates v_{rms}^2. This effect becomes smaller at higher redshifts, as shown in the other panels of Figure 3.2.

We find that both Equation (3.26) and (3.29) underestimate $P_{q_\perp}$ significantly compared to $P_{q_\perp}$ measured directly from the simulation. The underestimation decreases monotonically with redshift: $12-25\%$ over $k = 0.2-1.5\ h\ \mathrm{Mpc}^{-1}$ at $z=0$; $7-23\%$ at $z=1$; $5-18\%$ at $z=2$; and a few to 5% at $z=5.5$. Thus, this is likely related to the development of non-linear structure formation.

3.2.5.2 Connected Term

To show that the connected term is the likely explanation for the difference between the simulation data and the models that ignore the connected term, we calculate the connected term using perturbation theory. Since the connected term vanishes in linear theory, we must go to the next-to-leading order perturbation theory, such as the standard "one-loop" perturbation theory [7]. This theory allows us to extend validity of analytic solutions for Fourier modes of density and velocity fields down to weakly non-linear scales,tri e.g., in $k \lesssim 0.3\ h\ \mathrm{Mpc}^{-1}$ at $z=0$ and in wider wavenumbers at higher redshifts [36]. We derive the explicit expressions for the connected (Eq. 2.20) and unconnected terms (Eq. 2.27) in Appendix 2.2 and 2.3, and show them in the dotted and dashed lines in Figure 3.3, respectively.

We find that each of the unconnected and connected terms does not vanish in the low-k limit on its own. Instead, they converge to a constant value

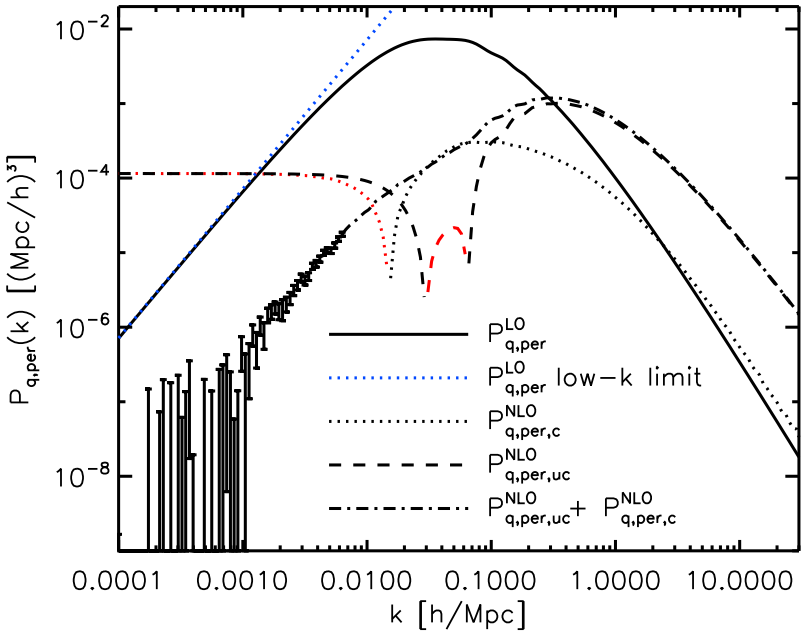

Figure 3.3: $P_{q_\perp}$ at $z = 0$ from the next-to-leading order unconnected term (dashed), connected term (black dotted), and the sum of the two (dot-dashed). The red color shows negative values. The error bars show the uncertainty of the Monte-Carlo integration for $k < 0.01\ h$ Mpc^{-1}. The solid line shows the LO term (Eq. 3.27), while the blue dotted line shows its low-k limit (Eq. 3.28).

with opposite signs. This constant can be obtained by taking $k \to 0$ limit of Equation (2.20) (Derived in Appendix 2.2.1):

$$\begin{aligned} P_{q_\perp,c}|_{k\to 0} &= \frac{16}{49}\int \frac{d^3k'}{(2\pi)^3}\int \frac{d^3k''}{(2\pi)^3}\frac{[1-(\hat{\mathbf{k}}\cdot\hat{\mathbf{k}'})(\hat{\mathbf{k}}\cdot\hat{\mathbf{k}''})]}{k'k''} \\ &\times P(\mathbf{k}')P(\mathbf{k}'+\mathbf{k}'')P(\mathbf{k}'')\left(\frac{(1-\mu^2)k''k'}{(\mathbf{k}'+\mathbf{k}'')^2}\right)^2. \end{aligned} \tag{3.32}$$

The sum of the two terms, however, yields a physical result that remains positive and decays toward lower k. The cancellation of two large numbers introduces an uncertainty in my numerical integration using the Monte Carlo method[3]. Within the uncertainty, the result is consistent with k^2 scaling, which is the same as the scaling of the LO expression given by Equation (3.28). This calls for a physical explanation; namely, what is the physical origin of the k^2 scaling in the low-k limit of the transverse momentum power spectrum, which is independent of cosmology or the initial power spectrum?

As shown by [44], this is a consequence of momentum conservation. In short, the argument goes as follows. Suppose that we have a uniform distribution of matter particles with no initial momentum or density fluctuation, and displace these particles with velocities $\mathbf{v}_i$ (where i is a particle ID) in a momentum-conserving manner. In this way, we have removed the effect of the initial condition, and can focus only on the effect of the subsequent evolution of particle's motion. Fourier transform of momentum of the displaced

[3]In principle we can derive the low-k limit of the next-to-leading order expression analytically. Here, we have chosen to perform numerical integration because of the complexity of the results given in Equations (2.20) and (2.27).

particles is given by $\mathbf{q}(\mathbf{k}) = \sum_i m_i \mathbf{v}_i \exp(-i\mathbf{k}\cdot x_i)$, and the low-$k$ limit is $\mathbf{q}(\mathbf{k}) = \sum_i m_i \mathbf{v}_i (1 - i\mathbf{k}\cdot\mathbf{x}_i + \dots)$. The first term vanishes by momentum conservation, giving $\mathbf{q}(\mathbf{k}) = \mathcal{O}(k)$; hence the power spectrum of momentum fields is proportional to k^2. This argument applies to both the longitudinal and transverse momentum fields. As we have found, the connected term is necessary for obtaining the correct low-k limit at the next-to-leading order in perturbation theory.

In the high-k regime, the connected term is of order ten percent of the unconnected term, which brings the predicted $P_{q_\perp}$ into better agreement with the simulation, as shown by the triangles in Figure 3.2. Thus, the underestimation is now much reduced to $2-18\%$ over $k = 0.2 - 1.5\ h\ \mathrm{Mpc}^{-1}$ at $z = 0$; $2-18\%$ at $z = 1$; $2-14\%$ at $z = 2$; and a few percent at $z = 5.5$. The remaining differences, especially at large k, are likely caused by inaccuracy of the next-to-leading order ("one-loop") calculation of perturbation theory, whereas the agreement at 2% level at small k supports my conclusion that the connected term is necessary for accurate modeling of the transverse momentum power spectrum.

3.2.6 Implication for the Post-reionization kSZ angular power spectrum

We use Equation (3.4) to calculate the observable CMB angular power spectrum of kSZ, C_ℓ. In the left panel of Figure 3.4, we show $dC_{\ell=3000}/ds$, which is the contribution to the kSZ signal at $\ell = 3000$ per comoving distance

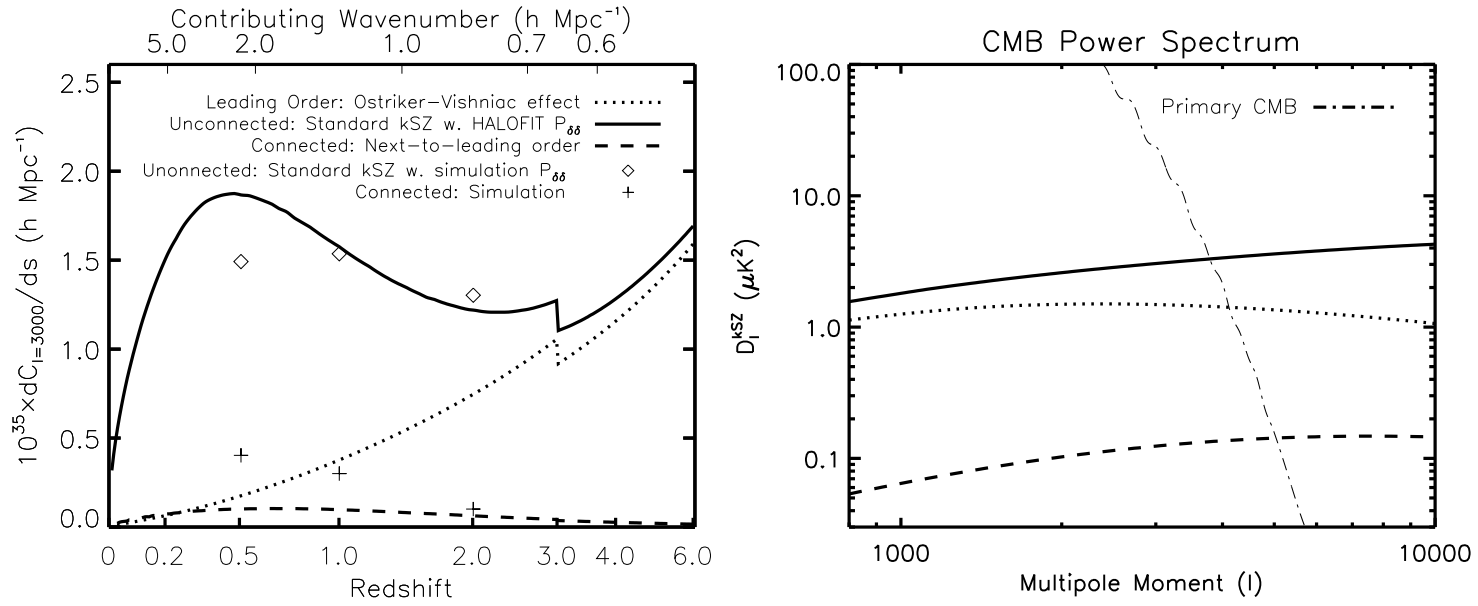

Figure 3.4: Angular power spectrum of the kSZ effect. (upper) $dC_{\ell=3000}/ds$, which shows the contribution to C_ℓ at $\ell = 3000$ from a given comoving distance. The horizontal axis shows redshifts such that it is linear in the comoving distance. The top label shows the contributing wavenumber to the kSZ signal, $\ell/s(z)$. The discontinuity at $z = 3$ is due to the instantaneous helium reionization. The solid line shows the Standard kSZ model (Eq. 3.29) with the HALOFIT density power spectrum, the dashed line shows the next-to-leading order connected term from perturbation theory, and the dotted line shows the LO contribution (Eq. 3.27). We also show the simulation results at $z = 0.5$, 1, and 2: the diamonds show Equation (3.29) with the non-linear density power spectrum from the simulation, while the crosses use the total simulation momentum power spectrum minus the diamonds. (lower) $D_\ell = \ell(\ell+1)C_\ell/(2\pi)$. The dash-dotted line shows the primary CMB temperature power spectrum, while the other lines show the same cases as in the upper panel.

from the observer. The area under each curve gives the total C_ℓ, which is shown in the right panel. The dotted line shows the LO calculation (Eq. 3.27). The solid line and the diamonds show the Standard kSZ model (Eq. 3.29) with the non-linear matter density power spectrum computed using the HALOFIT code [66] and the simulation (i.e., the square symbols in Figure 3.2), respectively. We cannot calculate the kSZ contribution at $z < 0.5$ in the simulation because the contributing wavenumber is beyond the resolution limit of the simulation ($k \sim 2\ h\ \mathrm{Mpc}^{-1}$). The disagreement between the two at $z = 0.5$ is due to the disagreement between the HALOFIT power spectrum and my simulation at $k \gtrsim 1\ h\ \mathrm{Mpc}^{-1}$, which is also due to the resolution limit of the simulation.

The dashed line shows the next-to-leading order connected moment contribution. We also estimate the connected term contribution in my simulation by subtracting the diamonds from the total transverse momentum power spectrum measured from the simulation. The simulation at $z = 1$ suggests about a factor of 3 larger effect than the next-to-leading order perturbation theory. The interpretation of the difference at $z = 0.5$ is difficult because of the disagreement between the diamond and the solid line; however, if we assume that the resolution of the simulation affects the unconnected and connected terms in the same way, then the simulation at $z = 0.5$ suggests about a factor of three larger effect than perturbation theory. This is expected because the kSZ signal receives contributions from $k \gtrsim 1\ h\ \mathrm{Mpc}^{-1}$ where the next-to-leading order calculation is insufficient to capture non-linearity (see the triangles in Figure 3.2). The difference decreases as z increases, as expected

from development of non-linear structure formation. Taking into account this extra non-linearity above perturbation theory, we estimate that the connected term contribution to C_ℓ at $\ell = 3000$ is at least ten percent of the unconnected term.

3.2.7 Summary and Conclusion

We have reexamined the currently popular model of the transverse dark matter momentum power spectrum (Eq. 3.29) used in the previous calculation of the post-reionization kSZ power spectrum [e.g., 64]. We find that the current model reproduces the contribution of the unconnected term (Eq. 3.26) well. However, this model ignores the contribution from the connected term that arises in the non-linear regime. Using both perturbation theory and the cosmological N-body simulation, we show that the contribution from the connected term adds a significant additional power, especially at larger k at lower z (Figure 3.2). This is consistent with the expectation from non-linear structure formation.

We estimate that the contribution of the connected term to the kSZ angular power spectrum at $\ell = 3000$ is ten percent of the unconnected term. This is the term that needs to be added to the calculation of [64], assuming that a similar correction is necessary for the momentum power spectrum of gas. In the light of the current observational constraint ($D^{\rm kSZ}_{\ell=3000} = 2.9\ \mu K^2$) and the post-reionization kSZ model of [64] ($D^{{\rm kSZ},z<6}_{\ell=3000} = 2.0\ \mu K^2$), adding ten percent to the post-reionization kSZ signal would imply twenty percent less

kSZ signal from the EoR. The semi-numerical model of [4] implies that such a change in the reionization kSZ signal would result in twenty percent increase in the reionization redshift, or ten percent shorter duration of reionization, Δz.

Finally, we have shown that both the LO and next-to-leading order perturbation theory calculations give $P_{q_\perp} \propto k^2$ in the low-k limit, independent of cosmology or the initial power spectrum. This behavior is consistent with momentum conservation.

3.3 The Kinetic Sunyaev-Zel'dovich effect as a probe of the physics of cosmic reionization: the effect of self-regulated reionization and sub-grid clumping model

[4]

Modeling the reionization contribution is a challenge, as the universe was not ionized homogeneously, but in patches. These patches grow over time until they overlap, finishing reionization of the universe. The distribution of these patches is determined by non-linear physics: non-linear clustering of the sources of ionizing photons; non-linear clumping of gas in the IGM; and complex morphologies of patches resulting from propagation of ionization fronts in the clumpy IGM. Accurately calculating the reionization contribution thus requires numerical simulations of cosmological structure formation coupled with radiative transfer.

[4]This chapter is based on a published work in [51].

To model the formation and spatial clustering of the sources of ionizing photons, cosmological simulations must be performed in a volume large enough to capture the crucial spatial variations of this process in a statistically meaningful way. This requires a volume greater than ~ 100 comoving Mpc across, because H II bubbles can typically grow as large as ~ 20 comoving Mpc in size. These simulations must also have a high enough mass resolution to resolve the formation of the individual galaxies which are the sources of ionizing radiation; thus, billions of particles are required. The radiative transfer of ionizing photons is then calculated on the IGM density and velocity fields computed by the cosmological simulation.

Following the early analytical calculation done for linear density and velocity perturbations in a fully ionized medium [48, 73, 35], calculations of the kSZ effect by cosmic reionization have steadily improved over time. Further analytical calculations later incorporated the effects of inhomogeneous reionization in an approximate manner [21, 59]. A "semi-numerical" approach was also developed by combining the simulated density and velocity fields from N-body simulations with an analytical ansatz for tracking the reionization process [78, 42].

Early, pioneering calculations using structure formation simulations coupled with radiative transfer to model inhomogeneous reionization numerically [19, 58] underestimated the amplitude of the kSZ signal, as they used computational boxes too small to capture the impact of large-scale velocity modes and H II bubbles or an accurate measure of the duration of the global

EOR. This was demonstrated by the first calculations of reionization based on truly large-scale (> 100 Mpc) radiative transfer simulations, which resolved the formation of all galactic halo sources above $2 \times 10^9\ M_\odot$ [31, 28]. These later simulations demonstrated the importance of a large enough simulation volume to capture the effects of long-wavelength fluctuations properly. They were also the first to realize that it is necessary to correct the kSZ power spectrum for the missing velocity power due to the finite box size of the simulations.

For the mass range of galactic halos resolved by these simulations, $\gtrsim 10^9\ M_\odot$, stars – the *sources* of reionization – were able to form when the primordial composition gas inside the halos cooled radiatively by atomic processes involving H atoms. They are known as "atomic cooling halos" to distinguish them from minihalos of mass $M \lesssim 10^8\ M_\odot$, with virial temperature $T_{\rm vir} \lesssim 10^4 K$, for which star formation is possible only if H_2 molecules form in sufficient abundance to cool the gas below $T_{\rm vir}$ by rotational-vibrational line excitation. Atomic-cooling halos with $10^8\ M_\odot \lesssim M \lesssim 10^9\ M_\odot$ also exist and are even more abundant than those with $M \gtrsim 10^9\ M_\odot$. These low-mass atomic-cooling halos ("LMACHs"), however, are prevented from forming stars if they form within an ionized patch of the IGM, where the gas pressure of the photoheated IGM opposes the accretion of baryons onto these halos. This "self-regulates" their contribution to reionization as the global ionized fraction grows with time and more and more of these halos are born within the ionized zones [60, 29]. While the precise value of halo mass which defines the upper edge of this "Jeans-filtered" mass-range is still uncertain, the high-mass

atomic-cooling halos ("HMACHs") above $\sim 10^9\ M_\odot$ are generally free of this suppression.

To simulate the impact of both LMACHs and HMACHs on reionization, it was necessary for [29] to increase their halo mass resolution so as to resolve all the LMACHs, too, by reducing the simulation box size to 53 Mpc on a side at fixed N-body particle number. This led to the first radiative transfer simulations of "self-regulated" reionization, which demonstrated the importance of including and then suppressing the LMACHs to start reionization earlier and extend its duration [29]. While the end of reionization is still set by the rapid rise of the HMACHs, in that case, when they eventually surpass the saturated contribution of the suppressible LMACHs, the effect of the LMACHs is to boost the electron-scattering optical depth, τ, integrated through the EOR. Such an effect can be important for the kSZ fluctuations from the EOR, too, but simulating this required us to increase the simulation volume again while retaining the high mass resolution required to resolve the LMACHs, too.

The next generation of simulations involved boxes 163 Mpc on a side, a volume large enough to predict observables like the kSZ effect, but with N-body simulations large enough to resolve all halos down to $10^8\ M_\odot$ and incorporate ionization suppression ("Jeans-filtering") of the halos of mass between $10^8\ M_\odot$ and $10^9\ M_\odot$ [30]. These smaller-mass halos (LMACHs) are more abundant and likely to be more efficient ionizing sources, as they may have higher escape fraction and emissivity [30]. However, as described above,

they may be suppressed as sources if they form inside ionized regions, where ionization heats the gas and makes its pressure high enough to resist gravitational collapse into such small galaxies. Recently, an additional simulation was performed, including this new physics, in an even larger volume (~ 600 Mpc) [27].

[1] expanded the mass range even further by accounting for starlight emitted by minihalos ($10^5 - 10^8\ M_\odot$), as well. In addition to their Jeans-mass filtering in ionized regions, they may also be suppressed if molecular hydrogen in minihalos is photo-dissociated by Lyman-Werner band photons in the UV background below 13.6 eV also emitted by the sources of reionization. They thus have a simulated model which takes into account all the halos down to $10^5\ M_\odot$ as sources of reionization.

It is important now to determine if and how the kSZ fluctuations from the epoch of reionization are different from the previous predictions when this "self-regulated" reionization is taken into account. That is the prime focus of this paper. Some of my results were first summarized in [61].

Recently, [45], [77] and [4] compared the predicted kSZ power spectra from their semi-numerical calculations, to the upper bounds from the SPT data [56], obtaining limits on the epoch and the duration of the reionization. Those studies concluded that, for a given value of the total Thomson-scattering optical depth, the reionization contribution to the kSZ signal is mostly sensitive to the duration of the reionization defined as $\Delta z \equiv z_{99\%} - z_{20\%}$ [77] or $z_{75\%} - z_{25\%}$ [45, 4]. [77] claim that the upper bound on $D^{\rm kSZ}_{l=3000}$ from the SPT data implies

$\Delta z < 4$ (95% CL) for no tSZ-CIB correlation, and $\Delta z < 7$ (95% CL) for the maximum possible tSZ-CIB correlation. However, as their methods are based on an analytical ansatz for the reionization process, it is necessary to use more self-consistent calculations of radiative transfer such as my simulation results to revisit this issue. We note that [77] compared their semi-numerical approach to their own numerical simulations using radiative transfer, finding an agreement at the level of 50%.

We shall also present kSZ predictions for models with varying assumption for the clumpiness of IGM. As mentioned in Chapter 1, spatially and time varying clumpiness of IGM will impact on the statistics in the ionization field, χ, and be reflected in the kSZ signal. While this is the next step of the work presented in Chapter 2, the EoR simulations used here is based on previous models and do not include my finding from Chapter 2, which is planned as a future work. The main analysis of these models will be shown in Mao et al. 2015 in preparation.

3.3.1 Basics

During the EoR, IGM is inhomogeneously ionized. We are thus required to account for spatial fluctuations in χ. We can incorporate this effect into

the equation by replacing δ in Equation (3.27) by $\chi(1+\delta)$:

$$\begin{aligned} & P_{q_\perp}(k,z) \\ = & \int \frac{d^3k'}{(2\pi)^3}(1-\mu'^2)\left[P_{\chi(1+\delta),\chi(1+\delta)}(|\mathbf{k}-\mathbf{k}'|)P_{vv}(k')\right. \\ & \left. -\frac{k'}{|\mathbf{k}-\mathbf{k}'|}P_{\chi(1+\delta),v}(|\mathbf{k}-\mathbf{k}'|)P_{\chi(1+\delta),v}(k')\right]. \end{aligned} \tag{3.33}$$

Note that we do *not* use this equation to compute $P_{q_\perp}$, but compute $P_{q_\perp}$ directly from the simulation. However, I use this equation to estimate and correct for the missing power due to a finite box size of the simulation as described in Chapter 3.3.1.1. We then use the corrected $P_{q_\perp}$ in Equation (3.4) to compute the angular power spectrum. As shown in the right panel of Figure 3.5, the effect of reionization inhomogeneity substantially boosts the power spectrum relative to the homogeneously-ionized case, while correcting for the missing velocity power of the finite simulation volume boosts it even further.

3.3.1.1 Correcting for the Missing Power in Simulations

The transverse momentum power spectrum at a given wavenumber, $P_{q_\perp}(k)$, receives contributions from the density and velocity auto/cross power spectra at various wavenumbers via Equation (3.33). As a result, $P_{q_\perp}$ computed from a simulation with a finite box suffers from a loss of power due to the lack of modes whose wavelength is greater than the size of the box [31].

The missing power arises because we do not have $P_{\chi(1+\delta),\chi(1+\delta)}(k)$, $P_{vv}(k)$, or $P_{\chi(1+\delta),v}(k)$ for $k < k_{\rm box} \equiv 2\pi/l_{\rm box}$, where $l_{\rm box}$ is the size of the

box. In Equation (3.33), this leads to the missing contributions in $|\mathbf{k}'| < k_{\rm box}$ and $|\mathbf{k}-\mathbf{k}'| < k_{\rm box}$. Estimating and correcting for the missing power thus requires the knowledge of the large-scale limit of $P_{\chi(1+\delta),\chi(1+\delta)}$, P_{vv}, and $P_{\chi(1+\delta),v}$.

For the homogeneous reionization case, it is straightforward to recover the missing power, as the large-scale limits of P_{vv}, $P_{\chi(1+\delta),\chi(1+\delta)}(=\bar{\chi}^2 P_{\delta\delta})$, and $P_{\chi(1+\delta),v}(=\bar{\chi}P_{\delta v})$ are precisely known by the cosmological linear perturbation theory. Using $P_{\delta\delta}$ from the linear theory and the linear relation, $P_{vv} = (\dot{a}f/k)^2 P_{\delta\delta}$, we find that the missing-power-corrected momentum power spectrum from the N-body simulation agrees precisely with the expected OV spectrum (see Figure 3.5). Note that most of the missing power comes from $P_{\delta\delta}(|\mathbf{k}-\mathbf{k}'|)P_{vv}(k')$ in $k' < k_{\rm box}$ because of the relation, $\mathbf{v}(\mathbf{k}) \propto \delta(\mathbf{k})/k$, in the large-scale limit.

For the inhomogeneous reionization case, we do not have a precise way to calculate the ionized density power, $P_{\chi(1+\delta),\chi(1+\delta)}$, in the large-scale limit; however, we expect that the density field and the ionization field are reasonably flat at the scales larger than the box size, and correct for the missing bulk velocity of the box. Therefore, we expect that the term $P_{\chi(1+\delta),\chi(1+\delta)}(|\mathbf{k}-\mathbf{k}'|)P_{vv}(k')$ in $k' < k_{\rm box}$ captures most of the missing power, as we have seen from the homogeneous reionization case above. With this approximation, the missing power in the inhomogeneously ionized regime is given by

$$P_{q_\perp}^{\rm Missing}(k,z) = \int_{k<k_{\rm box}} \frac{d^3k'}{(2\pi)^3}(1-\mu'^2)P_{\chi(1+\delta),\chi(1+\delta)}(|\mathbf{k}-\mathbf{k}'|)P_{vv}(k'). \quad (3.34)$$

In order to check the accuracy of Equation (3.34), we compare the missing-

power-corrected momentum power spectrum from the box of 114 h^{-1} Mpc (black solid line; denoted as L2) with that from a larger box of 425 h^{-1} Mpc (black dashed line; XL2) in Figure 3.8. We find a very good agreement between the two, confirming the robustness of my correction for the missing power.

3.3.2 Reionization Simulation

3.3.2.1 Basic simulation parameters

The simulations that we shall use in Chapter 3.3.2 consist of two parts: (1) cosmological N-body simulations of collisionless particles using the "CubeP^{3}M" N-body code [24]; and (2) radiative-transfer of H-ionizing photons in the density and source fields created from this N-body simulation results using the "C^2-Ray" (Conservative, Causal Ray-tracing) code [43]. The details of the simulations that we shall use in this paper are described in [30] and [1].

Unless specified otherwise, the reionization simulations are run on the density and source fields from the same N-body results with 3072^3 particles in a comoving box of 114 h^{-1} Mpc on a side. Halos are identified down to 10^8 $M_\odot$ with at least 20 particles, using a spherical overdensity halo finder with overdensity of 178 times the mean cosmic density. One of the models uses another N-body simulation with a larger box of 425 h^{-1} Mpc, with 5488^3 particles, resolving halos down to 10^9 $M_\odot$. The background cosmology is based on the *WMAP* 5-year data combined with constraints from baryonic acoustic oscillations and high-redshift Type Ia supernovae ($\Omega_M = 0.27, \Omega_\Lambda = 0.73, h = 0.7, \Omega_b = 0.044, \sigma_8 = 0.8, n_s = 0.96$; [38]).

For the 114 h^{-1} Mpc, the IGM density field from the particle data with halos excluded are calculated being adaptively-smoothed on to a 256^3 radiative-transfer grid in order to generate ionization maps using the C^2-Ray code. Therefore, the final physical length resolution of the reionization models is $d_{\rm cell} = 0.45\ h^{-1}$ Mpc. The highest l-mode that we can calculate from the simulation is given by $l_{\rm limit} = k_{\rm Nyq} s(z_{\rm ov})$, where $k_{\rm Nyq} = \pi/(2d_{\rm cell})$ is the Nyquist frequency, and $s(z_{\rm ov})$ is the comoving distance out to the end of reionization. For example, $z_{\rm ov} = 6.6$ gives $l_{\rm limit} = 22000$.

The new simulations also incorporate the effects of even smaller halos in $10^5\ M_\odot < M < 10^8\ M_\odot$, using a sub-grid prescription calibrated by smaller-box N-body simulations with higher-resolution having 1728^3 particles in a box of 6.3 h^{-1} Mpc [1]. Specifically, it is found that there is a correlation between the number of these small-mass halos in each cell and the total matter density averaged over that cell, with cells of size 0.45 h^{-1} Mpc, which coincides with the size of the radiative transfer cells in my 114 h^{-1} Mpc C^2-ray simulations. Then, this correlation is used to calculate the number of small-mass halos in each of the radiative-transfer cells in the 114 h^{-1} Mpc simulations.

For the most recent simulation, in a box 425 h^{-1} Mpc on a side, the RT grid has 504^3 cells, so $d_{\rm cell} = 0.84\ h^{-1}$ Mpc, slightly larger than that for the other simulations, and $l_{\rm limit} \sim 12000$. In this larger-box simulation, low-mass halos between 10^8 and $10^9\ M_\odot$ are included by a subgrid model like that described above for MHs.

3.3.3 Varying Properties of Ionizing Sources

Table 3.1: Reionization simulation parameters and global reionization history results

Label	Sources	$g_{\gamma,\mathrm{H}}$ ($f_{\gamma,\mathrm{H}}$)	$g_{\gamma,\mathrm{L}}$ ($f_{\gamma,\mathrm{L}}$)	$g_{\gamma,\mathrm{MH}}{}^{5}$ ($f_{\gamma,\mathrm{MH}}$)	τ_{es}	$z_{10\%}$	$z_{90\%}$	z_{ov}
L1	HMACHs+LMACHs	8.7(10)	130(150)	-	0.080	13.3	8.6	8.3
L2(XL2)	HMACHs+LMACHs	1.7(2)	8.7(10)	-	0.058	9.9	6.9	6.8
L2M1J1	HMACHs+LMACHs+MHs	1.7(2)	8.7(10)	5063(1030)	0.086	17.4	6.9	6.8
L3	HMACHs only	21.7(25)	-	-	0.070	10.3	9.1	8.4

What kind of sources are responsible for reionization? In this chapter, we consider a set of reionization simulations based on source models of increasing sophistication from the one with only high-mass sources to the one with all kinds of sources down to least massive halos in the models.

For each halo identified in the simulation, we calculate the number of ionizing photons which escape from it into the IGM per unit time, $\dot{N}_\gamma$, which is assumed to be proportional to the halo mass, M:

$$\dot{N}_\gamma = \frac{f_\gamma M \Omega_b}{\Delta t\, \Omega_0 m_p}\,, \tag{3.35}$$

where m_p is the proton mass, Δt is the duration of each star-forming episode (i.e. which corresponds in practice to the radiative transfer simulation time-step), and $f_\gamma = f_{\mathrm{esc}} f_\star N_\star$ is the number of ionizing photons produced and released by the halo over the lifetime of the stars which form inside it in this time step, per halo atom, if f_* is the fraction of the halo atoms which form stars during this burst, f_{esc} is the fraction of the ionizing photons produced by these stars which escapes into the IGM and the integrated number of ionizing

photons released over their lifetime per stellar atom is given by $N_\star$. The latter parameter depends on the assumed IMF for the stellar population and can range from $\sim 4,000$ (e.g. for Pop II stars with a Salpeter IMF) to $\sim 100,000$ (e.g. for a top-heavy IMF of Pop III stars). Halos were assigned different efficiencies according to their mass, grouped according to whether their mass was above ("HMACHs") or below ("LMACHs") $10^9 M_\odot$ (but above $10^8 M_\odot$, the minimum resolved halo mass). Low-mass sources are assumed to be suppressed within ionized regions (for ionization fraction higher than 10%), through Jeans-mass filtering, as discussed in [29].

In addition to the source efficiency parameter, f_γ, we also define a slightly different factor, g_γ, that is given by

$$g_\gamma = f_\gamma \left(\frac{10 \text{ Myr}}{\Delta t} \right) \tag{3.36}$$

where Δt is the time between two snapshots from the N-body simulation. The new factor g_γ reflects the fact that a given halo has a luminosity which depends on the ratio of f_γ to Δt, so g_γ has the advantage that it is independent of the length of the time interval between the density slices, and as such it allows a direct comparison between runs with different Δt. For the reader's convenience, we listed the values of both parameters in Table 3.1. The specific numerical values of the efficiency parameters are strongly dependent on the background cosmology adopted and the minimum source halo mass. Therefore, parameter values for simulations based on different underlying cosmology and halo mass resolution should not be compared directly, but require cosmology and

resolution-dependent conversion coefficients to achieve the same reionization history.

3.3.3.1 HMACHs-only model

In the simplest model (labeled as L3; see Table 3.1 for the details. Note that "L" stands for a "large volume"), only HMACHs are used as the sources of reionization. These sources are defined as the halos with $M > 2.2 \times 10^9 M_\odot$ for L3; and with $M > 10^9\ M_\odot$ (corresponding to $T_{\rm vir} \gtrsim 8 \times 10^4$ K at $z = 9$ from the TIS model of [33]) for the other configurations. These sources are believed to form stars even when immersed in ionized regions, due to the fact that their gravitational potential wells are deep enough to overcome Jeans-mass filtering.

3.3.3.2 HMACHs+LMACHs models

What about smaller-mass halos? LMACHs are more abundant; however, if they form inside the regions that have already been ionized, they would not act as sources of ionizing photons. This is because ionization heats the gas and makes its pressure too high for the gas to collapse into such small halos [29, and references therein].

When LMACHs are included to account for this "self-regulation" of reionization, LMACHs are given a higher efficiency, g_γ, than for HMACHs, as presumably it is easier for ionizing photons to escape from LMACHs than from HMACHs, and Pop III stars with a top-heavy initial mass function (IMF), which are capable of producing more ionizing photons than Pop II stars with

a Salpeter IMF, are more likely to form in LMACHs. If HMACHs are formed by mergers of smaller-mass halos, for example, they are more likely to have enough metallicity to make the transition from Pop III to Pop II star formation and, hence, to a less efficient IMF.

There are two cases which have both HMACHs and LMACHs, and we shall call them L1 and L2. For L1, the efficiency parameter, g_γ, is chosen such that the overlap redshift, $z_{\rm ov} = 8.3$, is similar to that of L3, $z_{\rm ov} = 8.4$ (see Table 3.1). For L2, g_γ is chosen such that $z_{\rm ov}$ is between 6 and 7, as suggested by the quasar absorption line observations.

For L2, we have another run with a much larger volume (425 h^{-1} Mpc) with 504^3 of radiative-transfer grids. Although it does not resolve LMACHs, LMACHs are included as a sub-grid model using correlation between average density of radiative transfer cells and number density of LMACHs similarly to how [1] included MHs in the simulation (Iliev et al. and Ahn et al. in preparation). This run gives $l_{\rm limit} \sim 12000$. We shall call this configuration "XL2", as the volume for this run is bigger (hence the name, XL) than those runs with "L." This run will be used to check my method to correct for the missing velocity power.

3.3.3.3 HMACHs+LMACHs+MHs model

What about *even* smaller-mass sources? Gas in halos of masses between $10^5\ M_\odot$ and $10^8\ M_\odot$ is thought to cool via rotational and vibrational transitions of hydrogen molecules and form stars, until hydrogen molecules are dissociated

by Lyman-Werner photons in the UV background from other sources [see [1] and references therein].

The MHs form earlier than LMACHs or HMACHs, and thus can start reionization of the universe earlier. However, as the star formation in MHs is vulnerable to Lyman-Werner photons, it gets suppressed wherever the intensity of the LW background rises above the threshold for suppression, locally at first, and eventually globally. This adds another kind of "self-regulation" to the reionization history, with an even more extended phase of low-level ionization before MHs are eventually suppressed completely [1].

The effects of MHs have been added to L2 by [1], and we take one of the cases simulated there, L2M1J1, as the fiducial case with MHs. See Table 3.1 for the efficiency of MHs. "M" denotes the mass spectrum of Pop III stars in MHs, and "J" the threshold intensity of the Lyman-Werner photon background, above which the star formation in MHs is suppressed. In L2M1J1, each halo is assumed to host one Pop III star with mass of 300 $M_{\odot}$, and the assumed LW threshold is $J_{\rm LW,th} = 10^{-22}$ ergs^{-1}cm^{-2}sr^{-1}.

This parameter choice for M_* and $J_{\rm LW,th}$ is only illustrative. As discussed in [1], the nature of the self-regulated suppression of MH star-formation is such that the contribution of MH stars to reionization rises to the point at which the global mean LW intensity reaches the threshold value for suppression. As long as MH stars dominate reionization (i.e. early phase), they continue to form at the global rate necessary to keep $J_{\rm LW}$ at this level, regardless of the value of M_*. For $M_* \gtrsim 100\ M_{\odot}$, the ratio of ionizing to dissociating

photons emitted per MH star is fixed, so their early contribution to reionization is also fixed by this self-regulation effect. Eventually, the LMACH and HMACH populations grow to dominate the LW background and suppress star formation inside MHs completely, thereafter. The value adopted for $J_{\rm LW,th}$ only affects the transition redshift at which this occurs somewhat (i.e. higher values allow MHs to contribute longer). In short, the reionization history is relatively insensitive to the value adopted for M_*, if $M_* \gtrsim 100\ M_\odot$, but somewhat more sensitive to $J_{\rm LW,th}$. Recent suggestion that MH stars may form with lower values of M_* (e.g. 40 $M_\odot$), perhaps with more than one star at a time, may alter some of these details, but the qualitative effect of self-regulation should remain. Similarly, the effect of a relative drift velocity between dark matter and baryons identified by [72], which tends to raise the minimum mass of MHs which typically form stars, is offset by a small shift in the timing of the early phase of MH-dominated reionization, as the exponential rise of MH abundances compensates at lower redshift.

3.3.4 Results and Conclusion : Varying Properties of Ionizing Sources

Before presenting and discussing my predictions for the kSZ power spectrum, let me briefly comment on the global ionization history of the universe, which is the key to understanding the difference between my results and the previous ones. For more detailed discussion on the effects of self-regulation, see [29, 30] and [1].

Figure 3.6 shows how the reionization proceeds in our simulation boxes,

while Figure 3.7 shows the mass-averaged ionization fraction of the universe as a function of redshift. Both figures show that inclusion of low-mass halos (LMACHs and MHs), which are self-regulated, significantly extends the ionization history of the universe toward higher redshift. Let me compare L1 and L3. As LMACHs form earlier, the universe begins to be ionized earlier in L1 than in L3. However, the universe does not get reionized quickly but keeps a low level of ionization for an extended period due to self-regulation of sources. Only after HMACHs start to dominate, at $z \sim 10$, does reionization proceed rapidly and finishes soon thereafter. In L3, with no LMACHs, by contrast, reionization proceeds rapidly from beginning to end because the abundance of HMACHs, the only sources, grows exponentially without any suppression effects to self-regulate them. When MHs are included (L2M1J2), the universe begins to be ionized even earlier than the cases with HMACHs and LMACHs, and keeps a low-level ionization for a longer period.

These physically motivated yet somewhat complex reionization histories were not considered in any of the previous calculations of the kSZ power spectrum. In this chapter, we show that it is these new features in the reionization history that invalidate simple two-parameter descriptions of the amplitude of the kSZ power spectrum proposed by the previous study [77, 45, 4].

3.3.4.1 Impact of Inhomogeneous Reionization

First, it is useful to understand how important it is to include inhomogeneity (or patchiness) of reionization when computing the kSZ power spec-

trum. In order to see this, we create a homogeneous version of L3 ("L3-homogeneous"), in which we wipe out inhomogeneity of reionization by replacing the ionization fraction, χ, with its global average, $\bar{\chi}$ (see Figure 3.7). This then gives the transverse momentum power spectrum as $P_{q_\perp} = \bar{\chi}^2 P^{\rm OV}_{q_\perp}$, where $P^{\rm OV}_{q_\perp}$ is the OV spectrum given by Equation (3.27). We remind reader that, on the scales of interest to me in this power spectrum ($k \lesssim 1\ h\ {\rm Mpc}^{-1}$), the degree of non-linearity of the underlying density and velocity fields of the IGM is small enough that we can well approximate the kSZ power spectrum for this "homogeneous" ionization case by the assumption of linear perturbations inherent in Equation (3.27) (see the left panel of Figure 3.5). We use this momentum power spectrum in Equation (3.4) to obtain the kSZ power spectrum for "L3-homogeneous." Thus, "L3" and "L3-homogeneous" have exactly the same average reionization history, while spatial fluctuations of ionization fraction are included only in L3. We find that L3 yields an order-of-magnitude larger power spectrum than L3-homogeneous that is consistent with findings in [31](see Figure 3.8).

In order to see the effect of inhomogeneous reionization on the kSZ power spectrum in more detail, we show the contribution from a given comoving distance to the kSZ power spectrum at $l = 3000$, $dC^{\rm kSZ}_{l=3000}/ds$, in Figure 3.9. While both L3 and L3-homogeneous converge to the same $dC^{\rm kSZ}_{l=3000}/ds$ after the universe becomes fully ionized, we find a clear enhancement of the power when the ionization fraction is less than unity, $z > z_{\rm ov} = 8.4$. The maximum contribution occurs when the universe is half ionized. One can see this

visually in the middle (L3) and bottom (L3-homogeneous) panels of Figure 3.9: L3 is clearly more patchy than L3-homogeneous.

The angular scale for $l = 3000$ roughly corresponds to the co-moving length of 15 h^{-1} Mpc during the reionization era ($z \sim 10$). The contribution to the kSZ power spectrum continues to grow until the typical comoving size of ionized bubbles reaches 15 h^{-1} Mpc. In the models, this occurs when the universe is half ionized. After this epoch bubbles grow bigger than 15 h^{-1} Mpc, and thus the ionization field is no longer patchy on the scale of 15 h^{-1} Mpc. This explains why the contribution to the kSZ power spectrum at $l = 3000$ decreases after the half-ionization epoch. (By the same token, a plot like that for the inhomogeneous case L3 in Figure 3.9 but for $l > 3000$ would look similar but with the peak shifted to higher z, when ionized patches were smaller-scale.)

3.3.4.2 Impact of LMACHs

How does the presence of LMACHs and self-regulation affect the kSZ power spectrum? To answer this we compare L1 and L3, which are mostly similar except that L1 has low-mass halos ($10^8\ M_\odot < M < 2.2 \times 10^9\ M_\odot$) with most of them being LMACHs. While they finish reionization at nearly the same redshift, L1 begins ionization earlier due to LMACHs and gives an extended period of low ionization due to self-regulation (see Figure 3.7).

Figure 3.8 shows that L1 and L3 give similar kSZ power spectra at $l \lesssim 3000$, while at higher multipoles L1 becomes significantly greater than

L3. This is because there are numerous ionized bubbles created by LMACHs at high redshifts, which give significant contributions to the small-scale kSZ power spectrum. Although it would be a challenge for current surveys, future measurements of $D_l^{\rm kSZ}$ with 10% accuracy over a wide range of multipoles can distinguish between the predictions of L1 and L3, shedding light on the roles of LMACHs during the reionization.

We compare the contributions from a given comoving distance to the kSZ power spectrum at $l = 3000$, $dC_{l=3000}^{\rm kSZ}/ds$, for L1 and L3 in Figure 3.10. As expected, L1 has larger contributions at higher redshifts ($z \gtrsim 10$) due to LMACHs. On the other hand, L3 has larger contributions at lower redshifts ($z \lesssim 10$), as it is more patchy due to the absence of smaller bubbles around LMACHs (see the middle (L3) and bottom (L1) panels of Figure 3.10). In L1, bubbles around LMACHs do not grow much because of self-regulation.

In the left panel of Figure 3.11, we show the cumulative contributions to the kSZ power spectrum at $l = 3000$ below a given maximum redshift, z. This also shows that L1 receives larger contributions from higher redshifts than L3: 20% of the total power in L1 comes from $z > 11$, while only a few percent of the total power in L3 comes from $z > 11$. Similarly, the right panel of Figure 3.11 shows that 20% of the total power in L1 comes from when the ionization fraction is less than 0.25, which is consistent with the ionization history above $z = 11$ shown in Figure 3.7. This extended tail has important implications for the interpretation of the kSZ power spectrum, as we shall discuss in Chapter 3.3.4.4.

3.3.4.3 Impact of Minihalos

What about MHs? We compare L2 and L2M1J1, which have the same efficiency parameters for HMACHs and LMACHs, but only L2M1J1 considering MHs. While L2 and L2M1J1 finish reionization at almost the same redshift, L2M1J1 begins ionization much earlier due to MHs and gives a significantly more extended period of low ionization due to self-regulation (see Figure 3.7).

Figure 3.8 shows that L2 and L2M1J1 give similar kSZ power spectra at $l \lesssim 5000$, while at higher multipoles L2M1J1 becomes greater than L2. The reason is the same as that for L1 versus L3: there are numerous ionized bubbles created by MHs at high redshifts, which contribute to the small-scale kSZ power spectrum.

While L2M1J1 begins reionization much earlier and thus has more contribution from high redshifts to the kSZ power spectrum, the actual magnitude of the high-redshift contribution is modest. This is because of self-regulation: self-regulation prevents bubbles around MHs from growing, and thus end up creating numerous small bubbles filling space nearly uniformly. This results in a lesser degree of patchiness, hence a modest contribution to the kSZ power spectrum at $l = 3000$. One can see this visually in the middle (L2M1J1) and bottom (L2) panels of Figure 3.12. As a result, the situation is similar to that for L1 versus L3: 20% of the total power at $l = 3000$ in L2M1J1 comes from $z > 10$, while only 5% of the total power in L2 comes from $z > 10$.

It is interesting that all the models with self-regulation (L1, L2, and L2M1J1) lie on top of each other when the cumulative contribution is shown as a function of the mean ionization fraction (see the right panel of Figure 3.11), whereas the model that does not have self-regulation (L3) is a clear outlier. Whether this is merely a coincidence or a unique feature of self-regulation is unclear due to the limited number of samples.

3.3.4.4 Spot checking the previous constraints on the duration of reionization: more extended histories can give similar kSZ signals

What determines the amplitude of the kSZ power spectrum? Recent studies using semi-numerical reionization models [77, 45, 4] claim that the amplitude of the kSZ power spectrum at $l = 3000$ can be described by a two-parameter family: the redshift of half-ionization, $z_{50\%}$, and the duration of reionization defined as $\Delta z \equiv z_{99\%} - z_{20\%}$ [77] or $\Delta z \equiv z_{75\%} - z_{25\%}$ [45, 4]. None of these studies included the effects of self-regulated reionization, and thus the reionization histories explored in these studies are roughly symmetric about the epoch of half-ionization.

Figure 2 of [77] shows that the kSZ power spectrum at $l = 3000$ increases by a factor of two as the duration of reionization increases from $\Delta z = 2$ to 4. Figure 10 of [45] shows that, for a half-ionization redshift of $z_{50\%} = 9$, the kSZ power spectrum at $l = 3000$ increases by a factor of 1.4 as the duration of reionization increases from $\Delta z = 1.3$ to 2.6. The former gives a scaling of $D^{\rm kSZ}_{l=3000} \propto (z_{99\%} - z_{20\%})$, whereas the latter gives $D^{\rm kSZ}_{l=3000} \propto (z_{75\%} - z_{25\%})^{0.5}$,

for a fixed half-ionization redshift. More recently, using a new semi-numerical method based on a correlation between the smoothed density field and the redshift-of-reionization field found from radiation-hydro simulations of [5], [4] calculate the kSZ power spectrum coming from $z > 5.5$ and obtain the following scaling relation:

$$D_{l=3000}^{\mathrm{kSZ,z>5.5}} = 2.02\ \mu\mathrm{K}^2 \left[\left(\frac{1+\bar{z}}{11}\right) - 0.12\right] \left(\frac{\Delta z}{1.05}\right)^{0.47}, \qquad (3.37)$$

where $\Delta z = z_{75\%} - z_{25\%}$ and $\bar{z}$ is the mean value of the redshift-of-reionization field, which is approximately equal to the half-ionization redshift, $z_{50\%}$.

Table 3.2: Global reionization history and kSZ signal

Label	$z_{50\%}$	$z_{99\%} - z_{20\%}$	$z_{75\%} - z_{25\%}$	z_{ov}	$D_{l=3000}^{\mathrm{kSZ,z>5.5}}$	$D_{l=3000}^{\mathrm{kSZ,z<z_{ov}\,6}}$	$D_{l=3000}^{\mathrm{kSZ,z>z_{ov}}}$	$D_{l=3000}^{\mathrm{kSZ,total}}$
L1	9.5	3.2	2.2	8.3	1.27	1.94	0.83	2.77
L2(XL2)	7.6	2.1	1.4	6.8	0.87	1.69	0.66	2.35
L2M1J1	7.7	6.5	2.1	6.8	0.90	1.69	0.69	2.38
L3	9.1	1.3	0.9	8.4	1.20	1.96	0.75	2.71

My predictions for $D_{l=3000}^{\mathrm{kSZ}}$ are summarized in Table 3.2. Among the models we have explored in this paper, L3 (which contains only HMACHs and does not have self-regulation) closely matches the scenarios explored in the above studies. Using $z_{50\%} = 9.1$ and $z_{75\%} - z_{25\%} = 0.9$ we find for L3, Equation (3.37) gives $D_{l=3000}^{\mathrm{kSZ,z>5.5}} = 1.5\ \mu\mathrm{K}^2$. This is in a reasonable agreement with my result,[7] $D_{l=3000}^{\mathrm{kSZ,z>5.5}} = 1.2\ \mu\mathrm{K}^2$.

[7] In order to compute $D_{l=3000}^{\mathrm{kSZ,z>5.5}}$, we calculate the contribution from z between 5.5 and z_{ov} using the fully-ionized formula, $P_{q_\perp} = P_{q_\perp}^{\mathrm{OV}}$, and add it to $D_{l=3000}^{\mathrm{kSZ,z>z_{ov}}}$ shown in the seventh column of Table 3.2.

However, the above formula significantly overestimates the amplitude of the kSZ power spectrum for L1: Equation (3.37) gives $D_{l=3000}^{\mathrm{kSZ,z>5.5}} = 2.4\ \mu\mathrm{K}^2$, whereas we find $D_{l=3000}^{\mathrm{kSZ,z>5.5}} = 1.3\ \mu\mathrm{K}^2$. In other words, despite the fact that L1 has a significantly more extended duration of reionization than L3 (by a factor of more than two), $z_{75\%} - z_{25\%} = 2.2$, the amplitude of the kSZ power spectrum increases only by 8%. Similarly, Equation (3.37) gives $D_{l=3000}^{\mathrm{kSZ,z>5.5}} = 1.5$ and $1.9\ \mu\mathrm{K}^2$ for L2 and L2M1J1, respectively, whereas we find $0.9\ \mu\mathrm{K}^2$ for both cases. Therefore, we conclude that Equation (3.37) is valid only for simple scenarios where the reionization history is roughly symmetric about the half-ionization redshift, but is invalid when self-regulation is included. Similar conclusions apply to [77] and [45].

My results show that self-regulation makes the duration of reionization significantly more extended without changing the amplitude of the kSZ power spectrum very much. In other words, an extended period of low-level ionization in $z > z_{50\%}$ does not make much contribution to the kSZ power spectrum at $l = 3000$.

3.3.4.5 Conclusion

In this dissertation, using the state-of-the-art reionization simulations incorporating the effects of self-regulated reionization [30, 1], In this dissertation, we have computed the power spectrum of the kSZ effect from the EOR using EoR simulations incorporating the effects of self-regulated reionization from [30, 1].

Unlike the previous work which created maps and computed two-dimensional Fourier transforms from the maps, we have computed the kSZ power spectrum from a line-of-sight integral of the transverse momentum power spectrum of ionized gas. We present a method to statistically correct for the missing velocity power in Chapter 3.3.1.1, and verify the accuracy of my method by comparing the results from large- (425 Mpc/h) and small-box (114 Mpc/h) simulations.

We find that the kSZ power spectrum is a sensitive probe of patchiness of reionization: patchiness increases the amplitude of the kSZ power spectrum by an order of magnitude. The maximum contribution occurs when the angular sizes of ionized bubbles are close to those corresponding to a given multipole.

While inclusion of small-mass halos such as LMACHs and MHs makes the beginning of reionization earlier, self-regulation significantly slows down the progress of reionization [29, 30, 1]. This results in an extended period of low-level ionization before more massive HMACHs dominate and finish reionization. We find that such an extended period of low-level ionization does not, however, make much of a contribution to the kSZ power spectrum at $l = 3000$: $D^{\rm kSZ}_{l=3000}$ changes only by $\sim 10\%$ despite the fact that the duration of reionization increases by a factor of more than two.

My results qualitatively change the conclusions reached by the previous work which did not include self-regulation. Recent work [77, 45, 4] assumes that $D^{\rm kSZ}_{l=3000}$ can be adequately parameterized by the redshift of half-ionization,

$z_{50\%}$, and the duration of reionization, Δz. While my result for the simplest model of reionization without self-regulation (L3) agrees with the scaling formula of [4] (Equation 3.37), my results for the models with self-regulation do not agree with it: specifically, the amplitude of the kSZ effect is no longer correlated well with the duration of the reionization. This is because self-regulation gives an extended period of low-level reionization only for $z > z_{20\%}$, while the simple models adopted by these other treatments have a roughly symmetric reionization history about $z = z_{50\%}$, for which a longer duration thus implies a longer period of patchy state with a significant ionization across $z = z_{50\%}$. Therefore, a more accurate scaling formula is required to take into account the asymmetric reionization history typical of self-regulated reionization.

Going beyond $l = 3000$, we find that LMACHs and MHs do have a considerable impact on the kSZ power spectrum on smaller angular scales. For example, $D^{\rm kSZ}_{l=10000}$ is boosted by 60% and 25% when LMACHs and MHs are included, respectively. Even though measurements of the kSZ power spectrum at $l > 3000$ would be a challenge for the moment due to contamination by extragalactic point sources and tSZ, future multi-wavelength observations may allow us to determine the kSZ power spectrum from the EOR over a wide range of multipoles. Such measurements will provide us with valuable additional information on the nature of the ionizing sources and the history of reionization.

How do my calculations compare with these current observational constraints? In order to obtain the total kSZ signal from both reionization

and post-reionization contributions, we take the "CSF" (cooling and star formation) post-reionization model of [64] that approximately incorporates the Jeans-filtering of $P_{q_\perp}$ due to shock heating in halos and in the IGM. The post-reionization kSZ signal computed from their scaling relation and the total kSZ signal (i.e., the sum of my reionization calculation and their post-reionization calculation) are shown in the sixth and seventh columns of Table 3.2, respectively. We find that all of my predictions are consistent with the 95% CL upper bound on the total signal from SPT, $D_{l=3000}^{\rm kSZ,total} < 2.8\ \mu K^2$ [56]. Therefore, we conclude that the current data are consistent with the current understanding of the physics of reionization.

3.3.5 Varying the Sub-grid Clumping model of the IGM

How clumpy was the IGM during the reionization. In this chapter, we consider a set of reionization simulations based on two clumping factor models as summarized in Table 3.3 and illustrated in Figure 3.13. Note that clumping factor for the cases in Table 3.1 is assumed to be unity. For the new models to be introduced here, the source property and efficiency are same as in L1. The two models are as follows.

(i) *inhomogeneous subgrid clumping* ("IC"): The clumping factor is calibrated from a high resolution N-body simulation with 6.3 $h^{-1}Mpc$ in side run with 3456^3 particles, which the mass is $5.12 \times 10^3 M_\odot$. They first grid this small-box simulation into 14^3 mesh, which the size is same as the RT mesh of 114 $h^{-1}Mpc$ simulation. The details of the calibration will be given in Mao et

Table 3.3: Reionization simulation parameters and global reionization history results for sub-grid Clumping models. All sub-grid clumping simulations are in a comoving volume of $114\,h^{-1}\,\mathrm{Mpc}$ on each side, which is coarse-grained onto a 256^3 mesh. The minimum mass source is halos with $10^8 M_\odot$, but halos with $10^8 M_\odot \leq M \leq 10^9 M_\odot$ are vulnerable to suppression if they formed inside an already ionized region. $z_{x\%}$ refer to the redshift when the ionized fraction reaches $x\%$. $z_{\rm ov}$ is the overlap redshift, i.e. when $\bar{x}_{\rm HII,m} = 0.99$.

label	acronym	subgrid clumping factor [a]	$\tau_{\rm es}$	$z_{10\%}$	$z_{50\%}$	$z_{75\%}$	$z_{90\%}$	$z_{\rm ov}$
"inhomogeneous clumping"	IC	$\hat{C}_{\rm cell}$ via fitting [b]	0.069	12.6	8.2	7.7	7.4	7.3
"biased homogeneous clumping"	BHC	$C_{\rm IGM,cell} = \overline{\hat{C}}$ [c]	0.067	12.6	8.0	7.3	7.1	6.9

[a] In all cases, the density that is multiplied by the clumping factor to compute the recombination rate is the inhomogeneous coarse-grained cellwise density in the RT mesh.

[b] The cellwise subgrid pseudo-clumping factor is interpolated from the local overdensity using equation (3.38) with best-fit coefficients in Table 3.4.

[c] The IGM subgrid clumping factor is everywhere equal to the global mean pseudo-clumping factor tabulated in Table 3.5.

al. 2015 in preparation.

To quantify this correlation, they use a polynomial fit

$$y = a_0 + a_1\, x + a_2\, x^2\,, \tag{3.38}$$

where $y = \log_{10} \hat{C}_{\rm cell}$ and $x = \log_{10}(1 + \langle\delta\rangle_{\rm cell})^2$, to fit to the scattered data at each redshift. Not only does the quadratic term make a second-order correction to the linear term, but it can characterize the concave nature in the correlation at low redshift. The coefficients a_0, a_1 and a_2 are redshift-dependent. Their best-fit values, using the least square method, are tabulated in Table 3.4.

(ii) *biased homogeneous subgrid clumping* ("BHC"): we improve upon models without sub-grid clumping factor, by allowing the subgrid clumping factor to increase with time in a one-zone approximation but still keep it

Table 3.4: The redshift dependent fitting of the cellwise pseudo-clumping factor $\hat{C}_{\rm cell}$ as a function of local overdensity $\langle\delta\rangle_{\rm cell} = \langle n_{\rm N,total}\rangle_{\rm cell}/\bar{n}_{\rm N,total} - 1$, $y = a_0 + a_1\,x + a_2\,x^2$, where $x = \log_{10}(1+\langle\delta\rangle_{\rm cell})^2$ and $y = \log_{10}\hat{C}_{\rm cell}$. The data is based on a coarse-grained mesh in which each cell is $0.45\,h^{-1}$cMpc on each side, using the $6.3\,h^{-1}$ Mpc N-body simulation which can resolve halos down to the Jean mass before reionization ($10^5\,M_\odot$).

z	a_0	a_1	a_2	z	a_0	a_1	a_2
60.000	0.00124	0.0463	0.0594	11.090	0.733	0.519	-0.200
41.106	0.0257	0.0802	0.0899	10.877	0.748	0.507	-0.207
38.919	0.0307	0.0885	0.100	10.673	0.763	0.498	-0.203
36.996	0.0358	0.0977	0.112	10.478	0.780	0.497	-0.203
35.289	0.0411	0.108	0.128	10.290	0.795	0.485	-0.217
33.761	0.0466	0.120	0.145	10.110	0.807	0.480	-0.198
32.385	0.0524	0.132	0.167	9.938	0.823	0.469	-0.213
31.137	0.0584	0.147	0.190	9.771	0.834	0.472	-0.191
30.000	0.0647	0.163	0.216	9.611	0.851	0.460	-0.212
27.900	0.0795	0.202	0.269	9.457	0.863	0.452	-0.220
26.124	0.0964	0.247	0.319	9.308	0.875	0.441	-0.225
24.597	0.116	0.297	0.349	9.164	0.887	0.446	-0.199
23.268	0.138	0.347	0.359	9.026	0.897	0.440	-0.193
22.100	0.163	0.395	0.350	8.892	0.908	0.429	-0.197
21.062	0.190	0.440	0.320	8.762	0.918	0.423	-0.193
20.134	0.219	0.481	0.283	8.636	0.931	0.418	-0.211
19.298	0.249	0.513	0.224	8.515	0.944	0.412	-0.215
18.540	0.279	0.539	0.162	8.397	0.952	0.399	-0.208
17.848	0.310	0.563	0.114	8.283	0.962	0.401	-0.202
17.215	0.341	0.579	0.0623	8.172	0.973	0.394	-0.206
16.633	0.371	0.588	0.0147	8.064	0.981	0.387	-0.202
16.095	0.400	0.594	-0.0191	7.960	0.989	0.378	-0.205
15.596	0.428	0.598	-0.0472	7.859	0.999	0.380	-0.195
15.132	0.456	0.596	-0.0890	7.760	1.005	0.364	-0.200
14.699	0.486	0.601	-0.107	7.664	1.014	0.363	-0.203
14.294	0.509	0.588	-0.151	7.570	1.024	0.356	-0.211
13.914	0.533	0.586	-0.159	7.480	1.035	0.362	-0.204
13.557	0.557	0.581	-0.173	7.391	1.042	0.350	-0.217
13.221	0.580	0.581	-0.172	7.305	1.047	0.348	-0.204
12.903	0.602	0.567	-0.192	7.221	1.057	0.344	-0.206
12.603	0.623	0.558	-0.205	7.139	1.065	0.334	-0.212
12.318	0.642	0.558	-0.183	7.059	1.070	0.334	-0.197
12.048	0.662	0.558	-0.198	6.981	1.078	0.328	-0.198
11.791	0.680	0.539	-0.205	6.905	1.086	0.328	-0.201
11.546	0.698	0.533	-0.202	6.830	1.094	0.317	-0.203
11.313	0.715	0.524	-0.195	6.757	1.103	0.316	-0.202

homogeneous. We use the global mean pseudo-clumping factor $\overline{\hat{C}}$ defined as

$$\overline{\hat{C}} \equiv \overline{n^2}_{\rm N,IGM}/\bar{n}^2_{\rm N,total}\,, \tag{3.39}$$

where $\overline{n^2}_{\rm N,IGM}$ is the global average of the IGM quadratic density field $\left\langle n^2_{\rm N,IGM}\right\rangle_{\rm cell}$ (note that $\overline{\hat{C}}$ is not the volume-weighted mean of $\hat{C}_{\rm cell}$).

For the small box (6.3 h^{-1}Mpc) N-body simulation, they smooth the full N-body particle data with SPH-like adaptive kernel to compute the cellwise $\left\langle n^2_{\rm N,IGM}\right\rangle_{\rm cell}$ using

$$\left\langle n^2_{\rm N,IGM}\right\rangle_{\rm cell} \equiv \frac{1}{V_{\rm cell}}\int_{\rm cell} n^2_{\rm N,IGM}(\mathbf{r})d^3r \tag{3.40}$$

$$= \frac{1}{V_{\rm cell}}\sum_{i\in{\rm IGM}} n_{i,{\rm IGM}}\int_{\rm cell} W(\mathbf{r}-\mathbf{r}_i;h_i)d^3r\,, \tag{3.41}$$

average it over the whole simulation volume to get $\overline{n^2}_{\rm N,IGM}$, and compute $\overline{\hat{C}}$. Note that this approach is equivalent to directly smoothing over all N-body particle data in the simulation volume, so the values of $\overline{\hat{C}}(z)$ depend only on the N-body fine cell resolution (e.g. $2^3\,N_{\rm particle} = 3456^3$ fine cells for the 6.3 h^{-1}Mpc N-body simulation), not on the coarse-grained mesh resolution. We tabulate the results in Table 3.5.

On the other hand, for the large box (114 h^{-1}Mpc) simulation where the gas clumping is poorly resolved, we use the small box (6.3 h^{-1}Mpc) result as the subgrid recipe as follows. We obtain the local clumping factor $\hat{C}_{\rm cell}$, for each cell of the large box coarse-grained mesh, from the coarse-grained mesh data of overdensity using the fitting formula (equation 3.38 and Table 3.4),

compute the cellwise $\left\langle n^2_{\rm N,IGM}\right\rangle_{\rm cell} = \hat{C}_{\rm cell}\left\langle n_{\rm N,total}\right\rangle^2_{\rm cell}$, and then average it to compute $\overline{n^2}_{\rm N,IGM}$ and $\overline{\hat{C}}$, which is tabulated in Table 3.5).

3.3.6 Results and Conclusion: Varying the Sub-grid Clumping model of the IGM

We compare the kSZ power spectra from L1 (red dotted), L2 (green long-dashed), IC (black solid) and BHC (blue dashed) in Figure 3.14 (upper panel). The positive slope of the kSZ power spectrum at $l \lesssim 3000$ reflects the geometry of reionization. If the reionization is dominated by smaller H II regions, the resulting slope is higher [51]. We find that the IC model yields the larger slope of the kSZ power spectrum than the L2 model, which reflects the fact that the IC model accommodates a larger population of smaller ionized bubbles at a given ionized fraction. Also, the kSZ power spectra of the BHC and IC models have almost identical slopes, which implies that their characteristic sizes of H II regions are close, though not identical, to each other.

The lower panel of Figure 3.14 shows the contribution of the kSZ power spectrum at $l = 3000$ from different redshifts. For a given clumping model and its reionization simulation, the kSZ power spectrum is dominated by the contribution from a narrow range of redshifts. It is interesting to see that the peak of this distribution appears when $\bar{x}_{\rm HII,m} \gtrsim 0.5$ for all reionization models considered, so it reflects the global history of reionization: the faster the reionization proceeds, the earlier the peak contribution of the kSZ power spectrum appears. We find that the redshift of this peak follows exactly the

Table 3.5: The global mean pseudo-clumping factor $\overline{\hat{C}}(z) \equiv \overline{n^2}_{\rm N,IGM}/\bar{n}^2_{\rm N,total}$ as a function of redshift z, calculated using the $6.3\,h^{-1}\,$Mpc simulation data.

z	$\overline{\hat{C}}$	z	$\overline{\hat{C}}$
60.000	1.009	10.478	9.053
41.106	1.077	10.290	9.302
38.919	1.092	10.110	9.938
36.996	1.108	9.938	9.996
35.289	1.125	9.771	10.67
33.761	1.143	9.611	10.81
32.385	1.163	9.457	11.12
31.137	1.184	9.308	11.43
30.000	1.207	9.164	12.33
27.900	1.264	9.026	12.80
26.124	1.335	8.892	12.98
24.597	1.422	8.762	13.50
23.268	1.526	8.515	13.88
22.100	1.648	8.397	14.27
21.062	1.789	8.283	14.93
20.134	1.949	8.172	15.27
19.298	2.122	8.064	15.64
18.540	2.310	7.960	15.82
17.848	2.521	7.859	16.58
17.215	2.740	7.760	16.58
16.633	2.971	7.664	17.20
16.095	3.217	7.570	17.11
15.596	3.476	7.480	19.02
15.132	3.725	7.391	17.95
14.699	4.060	7.221	19.07
14.294	4.251	7.139	19.07
13.914	4.566	7.059	20.04
13.557	4.866	6.981	20.35
13.221	5.220	6.830	21.02
12.903	5.488	6.757	21.58
12.603	5.761	6.686	21.99
12.318	6.211	6.617	22.48
12.048	6.511	6.549	22.02
11.791	6.805	6.483	21.96
11.546	7.222		
11.313	7.580		
11.090	7.946		
10.877	8.187		

order in $z_{\rm ov}$ i.e. (from the earliest to the latest arrival) the L1, IC, BHC, and L2 model.

The amplitude of the kSZ power spectrum in Figure 3.14 (upper panel) depends both on the peak location and on the amplitude of the peak. If the peak amplitude is fixed, then the earlier the peak appears, the larger the total amplitude of the kSZ power spectrum is. If the peak redshift is fixed, then a larger peak amplitude surely enhances the total kSZ amplitude. We find that the L1 model has the largest total amplitude, the IC model the second, and the BHC and L2 models the smallest[8], which is consistent with the locations of peak for different models. Note that the amplitude of the kSZ power spectrum in the BHC model is about 10% (relative error) smaller than in the IC model.

[8]Careful readers may find that the amplitude of the kSZ power spectrum in the L2 model is larger than in the BHC model for $l \lesssim 3000$, while reionization proceeds slightly slower in the former. This is because the amplitude of the kSZ peak contribution in the BHC model is smaller.

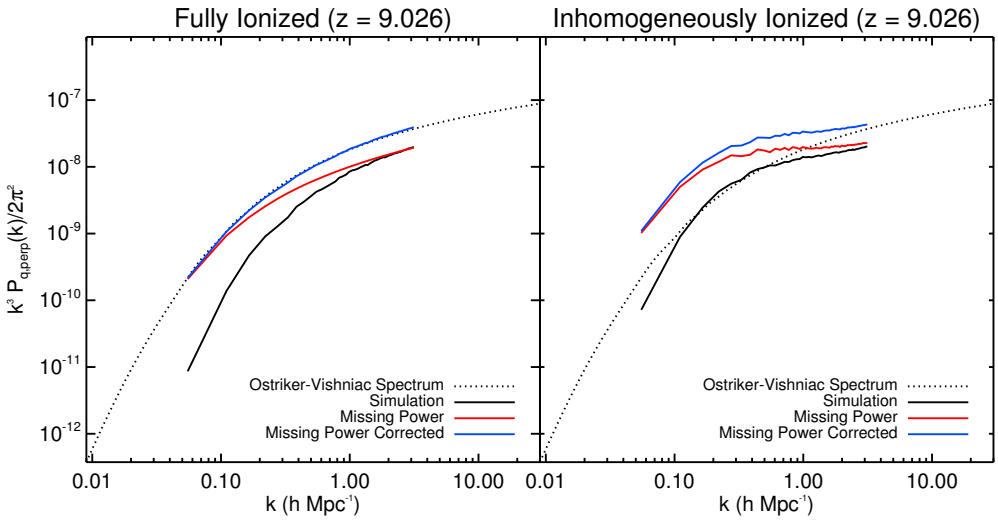

Figure 3.5: Dimensionless power spectra of the curl of the momentum field, $k^3 P_{q_\perp}(k)/(2\pi^2)$, at $z = 9$ calculated from the simulation with 114 h^{-1} Mpc in a side. The black solid lines show the raw power spectrum obtained from the N-body simulation, while the blue lines show the power spectrum after corrected being for the missing velocity power due to a finite box size of the simulation. The red lines show the missing power added to the black solid lines. The dotted lines show the analytical OV spectrum given in Equation (3.27). Left: fully ionized case. An excellent agreement between the OV spectrum and the corrected power spectrum shows the validity of the simulation as well as that of my method to correct for the missing velocity power. Right: inhomogeneously ionized case, L3. The power spectrum is significantly enhanced at $k \lesssim 1\ h$ Mpc^{-1}.

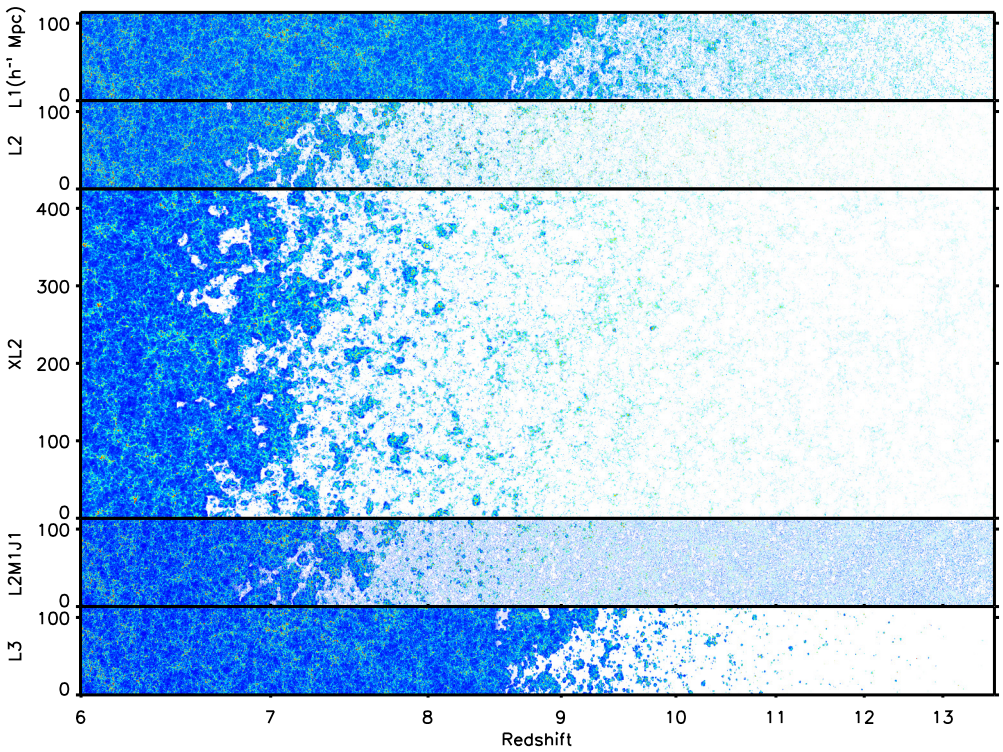

Figure 3.6: Cuts through the N-body+Radiative Transfer simulations used in this work. See Table 3.1 for the parameters of models L1, L2, L2M1J1, and L3. While these runs have the box size of 114 h^{-1} Mpc, the model XL2 has the box size of 425 h^{-1} Mpc and has the same model parameters as the model L2. Each panel shows the matter density distribution *multiplied by spatially-varying ionization fractions*. For example, it just shows the matter density when a given region is fully ionized, while it shows nothing (i.e., white) when a given region is fully neutral. The density fields are color-coded such that overdense regions are red and underdense regions are blue. We create this figure by interpolating between adjacent snapshots at a given lookback time. The length scale is linear in the co-moving units. The x-axis shows redshifts, while the y-axis shows h^{-1} Mpc.

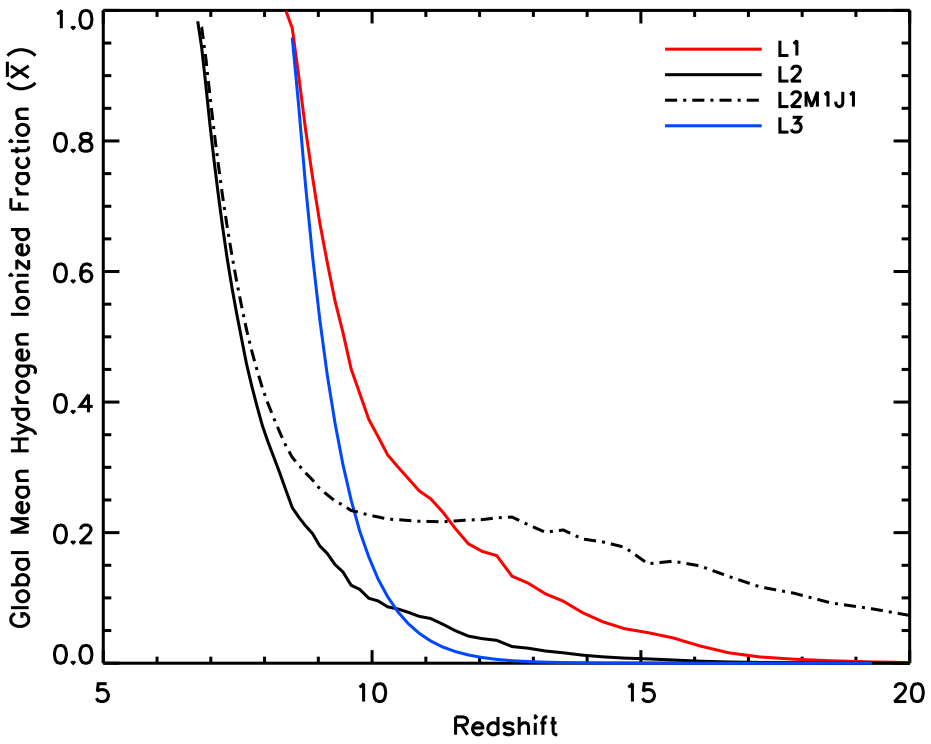

Figure 3.7: The global mean ionization history of the models (see Table 3.1 for the parameters of models). The mass-averaged hydrogen ionization fraction, $\bar{X}$, is plotted against z. Note how self-regulation results in an extended period of low-level ionization by comparing the case without self-regulation (L3 = HMACHs only) and that with self-regulation (L1 = HMACHs + LMACHs) [30]. A further extension occurs when MH sources are included, as well (i.e. compare L2 = HMACHs + LMACHs and L2M1J1 = L2 + MHs) [1].

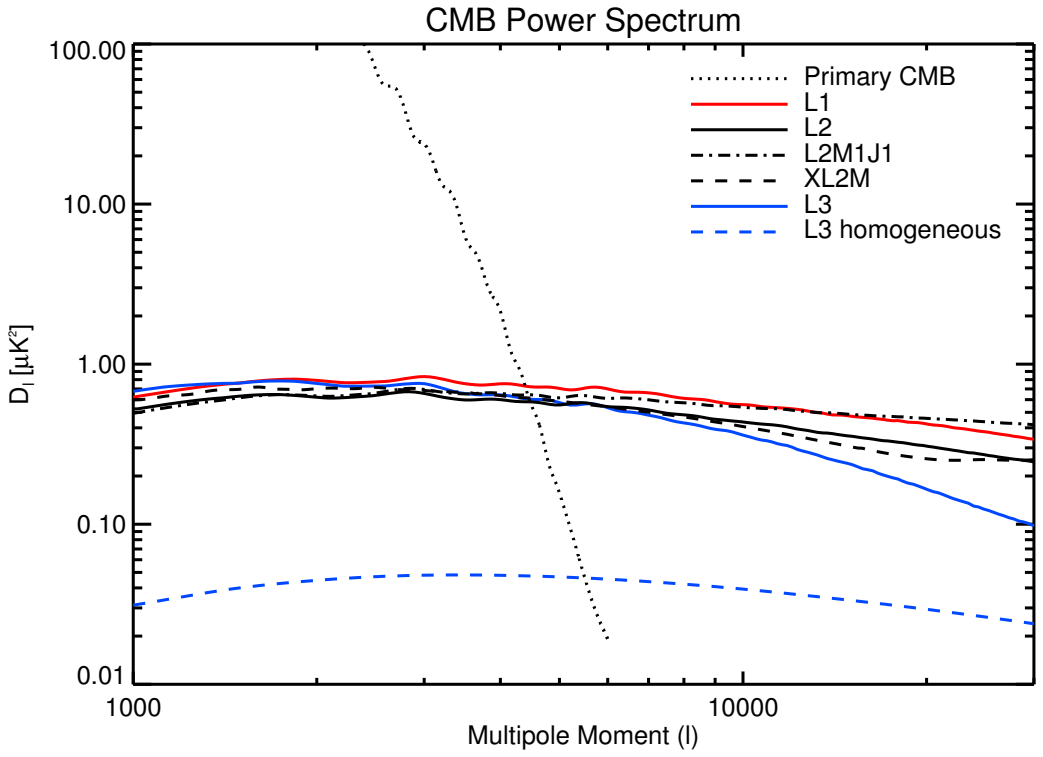

Figure 3.8: Predicted kSZ power spectra, $D_l^{\rm kSZ}$, from $z > z_{\rm ov}$ for the models discussed in this work (see Table 3.1 for the parameters of models). $z_{\rm ov}$ = 8.3, 6.8, 6.8, 6.8 and 8.4 for L1, L2, XL2, L2M1J1 and L3, respectively. The box size of L1, L2, L2M1J1 and L3 is 114 h^{-1} Mpc, while that of XL2 is 425 h^{-1} Mpc. The model parameters of XL2 are the same as those of L2, and thus XL2 provides a useful check of the way we correct for the missing velocity power in 114 h^{-1} Mpc-box simulations (see Chapter 3.3.1.1 for details). The primary CMB power spectrum is also shown.

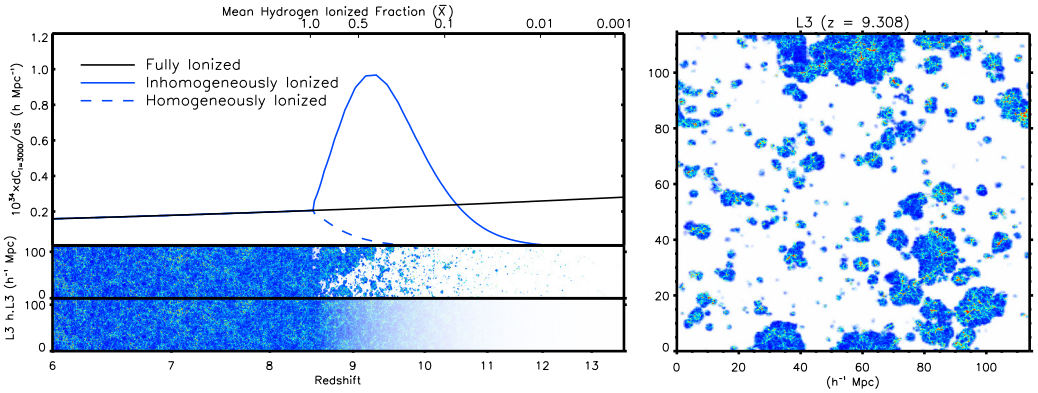

Figure 3.9: Left: The top panel shows the contribution from a given comoving distance to the kSZ power spectrum at $l = 3000$, $dC^{\rm kSZ}_{l=3000}/ds$. The solid line with a peak shows L3, the dashed line shows L3-homogeneous, and the nearly-horizontal solid line shows the fully-ionized case. The middle panel is the same as the bottom panel of Figure 3.6. The bottom panel shows L3-homogeneous, i.e., the density distribution multiplied by the average ionization fraction. Right: A snapshot of L3 at $z = 9.3$, which gives the maximum contribution to the kSZ power spectrum at $l = 3000$.

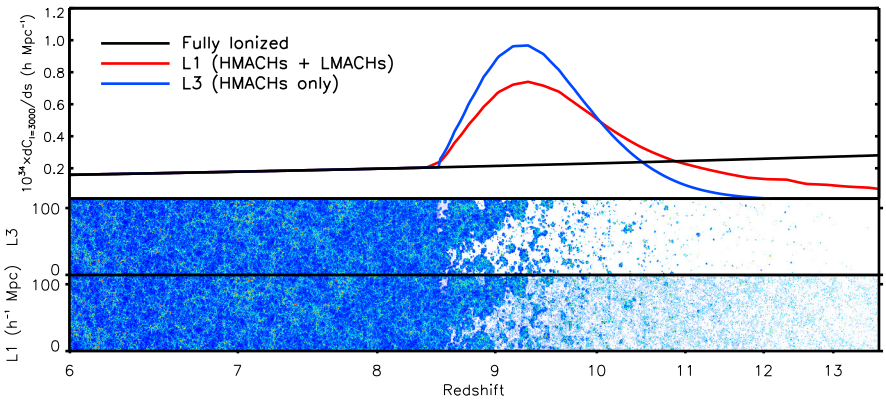

Figure 3.10: Same as the left panel of Figure 3.9, but for comparing L1 (bottom panel) and L3 (middle panel). See Table 3.1 for the parameters of L1 and L3.

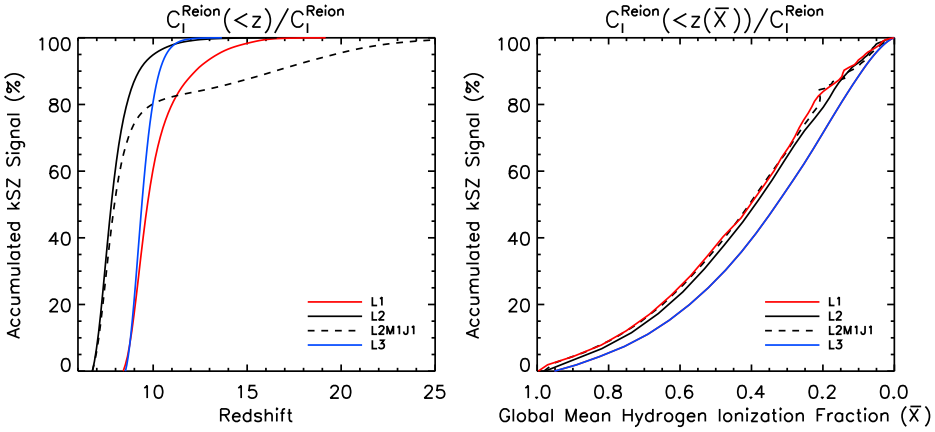

Figure 3.11: Cumulative reionization kSZ power spectrum at $l = 3000$ as a function of the maximum redshift (Left) and the mean ionization fraction (Right).

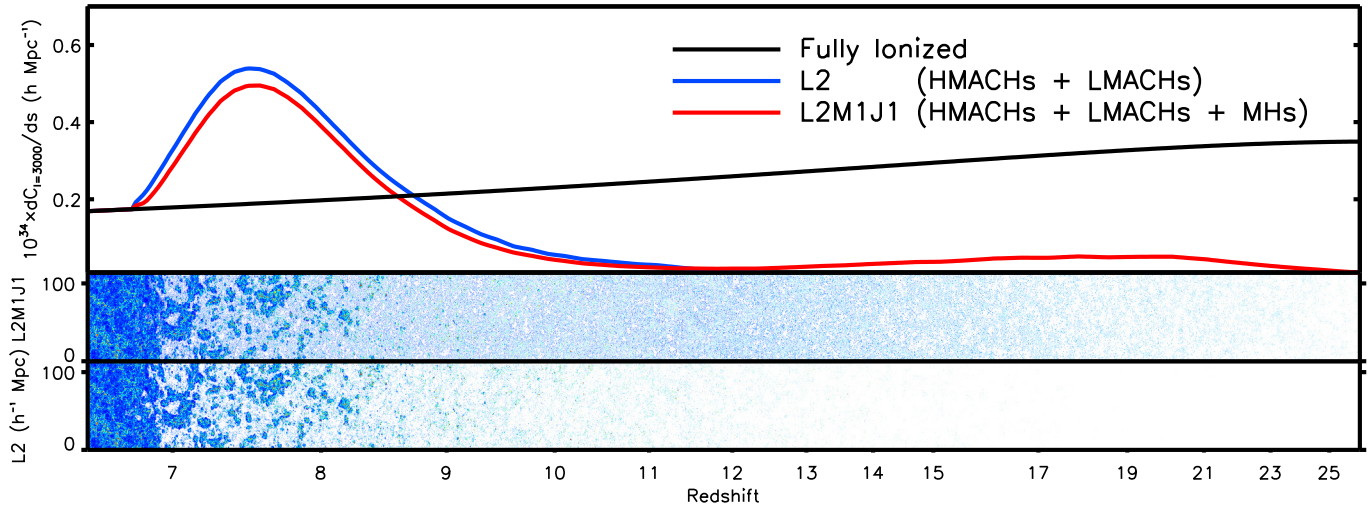

Figure 3.12: Same as the left panel of Figure 3.9, but for comparing L2 (bottom panel) and L2M1J1 (middle panel). See Table 3.1 for the parameters of L2 and L2M1J1.

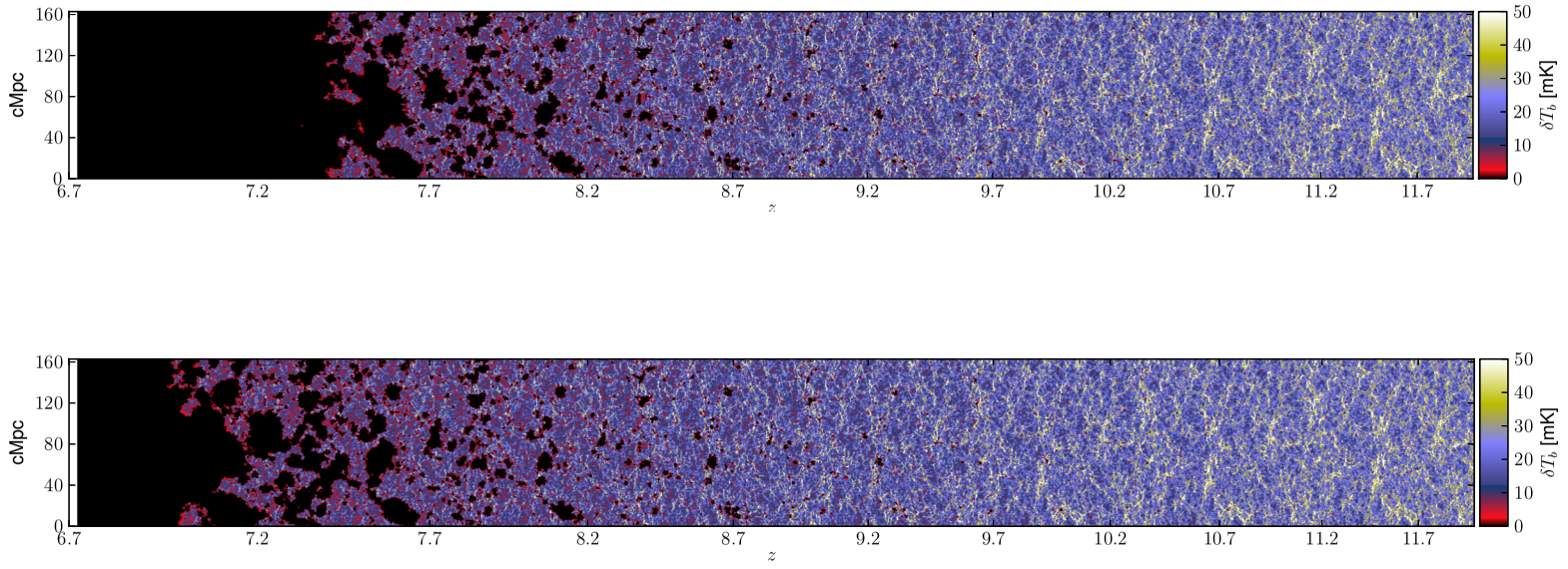

Figure 3.13: Position-redshift slices from inhomogeneous clumping case and biased homogeneous clumping case. These slices illustrate the large-scale geometry of reionization and the significant local variations in reionization history as seen at redshifted 21 cm line. Observationally they correspond to slices through an image-frequency volume of a radio array. The images show the differential brightness temperature at the full grid resolution in linear scale. The spatial scale is given in comoving Mpc. We note that for visualization purposes we artificially set $x_{\rm HI} = 10^{-5}$ after reionization ($z < z_{\rm ov}$). The redshift-space distortions due to the peculiar velocities are also included. Figure from Mao et al. (2015) in preparation.

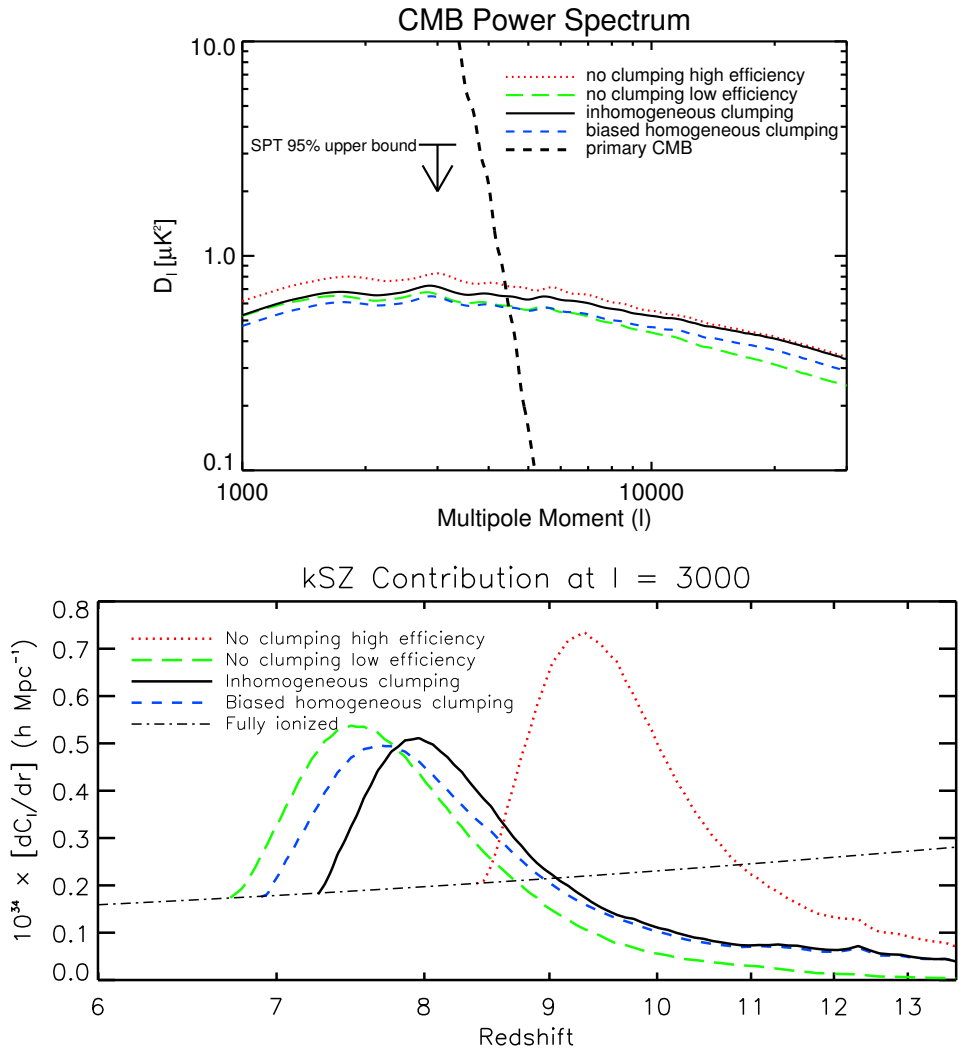

Figure 3.14: (upper) The kinetic Sunyaev-Zel'dovich effect from $z > z_{\rm ov}$ (see Table 3.3 for the value of $z_{\rm ov}$ for each model). The power spectrum of the primary CMB is shown as the thick dashed line for comparison. The 95% upper bound of $D_{l=3000}$ from the South Pole Telescope measurement [18] subtracted by the post-reionization kSZ signal from the cooling and star-formation model of [64] that is re-scaled to our cosmology is shown as a downward arrow. (lower) The history of contribution of the kSZ signal at $l = 3000$ in terms of the contribution per comoving distance, dC_l/dr. The nearly-horizontal dot-dashed line shows the case that assumes all the gas is ionized. Figure from Mao et al. (2015) in preparation.

Appendix

Appendix 1

Test Problem : Evaporation of a Spherical Halo

We test the accuracy of the GADGET-RT code for a spherically symmetric configuration that the one-dimensional radiation-hydrodynamics code of [2] can reproduce. In their work, the 1D code was used to assess the effect of EIBR on a minihalo with minimum-energy truncated isothermal sphere (TIS) profile [62, 33]. The 1D code accurately captures the evolution of ionization fronts both in the supersonic r-type phase and the subsonic d-type phase. The 1D code has been tested for a number of problems with analytical solutions (See Appendix C of their work).

We, first, create an initial condition for the 1D code using the fitting formula in Appendix A of [62]. We adopt $M = 10^6\ M_\odot$ for the mass inside the truncation radius, $r_t = 170\ physical\ pc$, and $z_{\rm col} = 10$ for the redshift of collapse. We do *not* actually truncate the initial condition at $r = r_t$, but extend it to $r = 10r_t$ using the fitting formula in order to test the code for the outskirt of the halo. This extended TIS profile decays toward large r in the reasonable way. The density profile at $\Delta = 1.1$ Myr shown as black lines in Figure 1.2 preserves the initial condition. We keep track of $N_{\rm sh}$=10,000

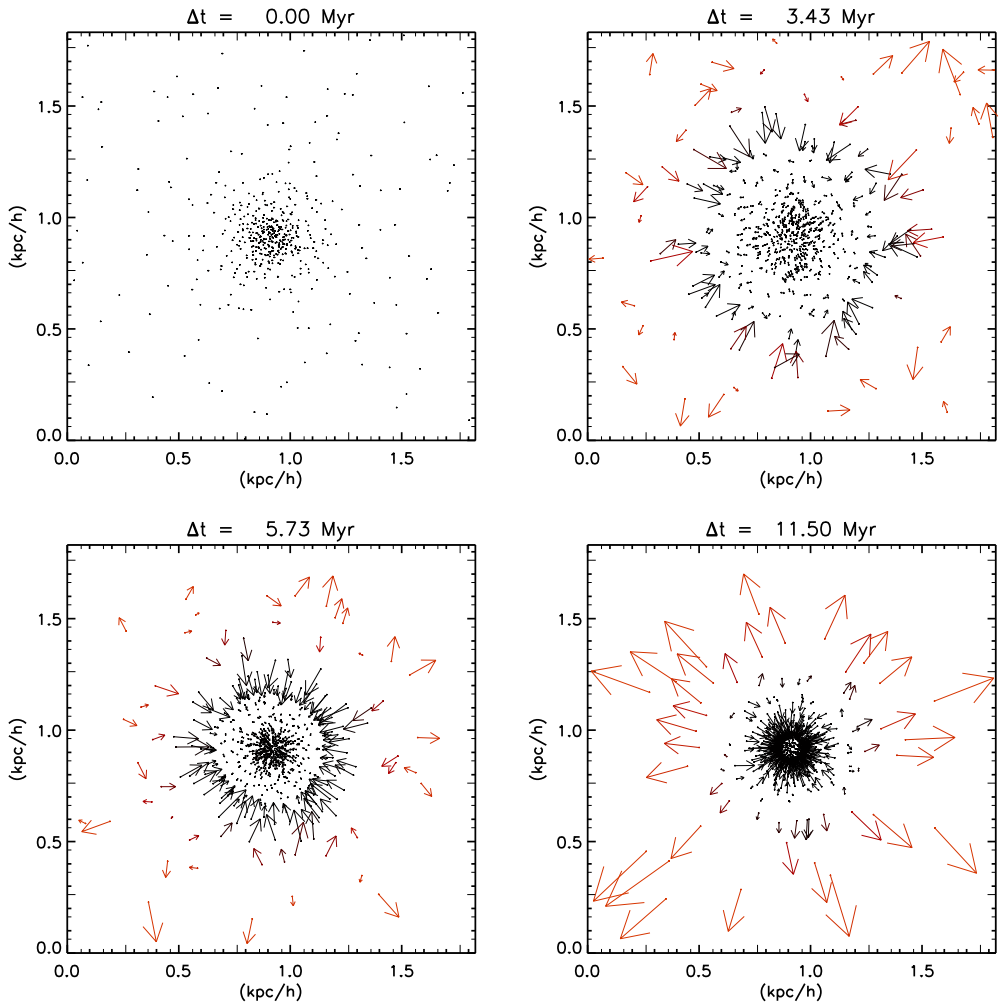

Figure 1.1: SPH particle map of the GADGET-RT simulation at Δt =0 (upper left), 3.43 (upper right), 5.73 (lower left) and 11.5 Myr (lower right). The map samples particles in a plain that goes through the center of the halo and the thickness is 0.2% of the simulation box (18.2 h kpc comoving). Only 20% of the particles are displayed for visual convenience. The arrows denote the projected particle velocities with the positions of their heads giving the linearly extrapolated positions after 5 Myr. Color is used to describe the ionization status of the particle (red-ionized, black-neutral).

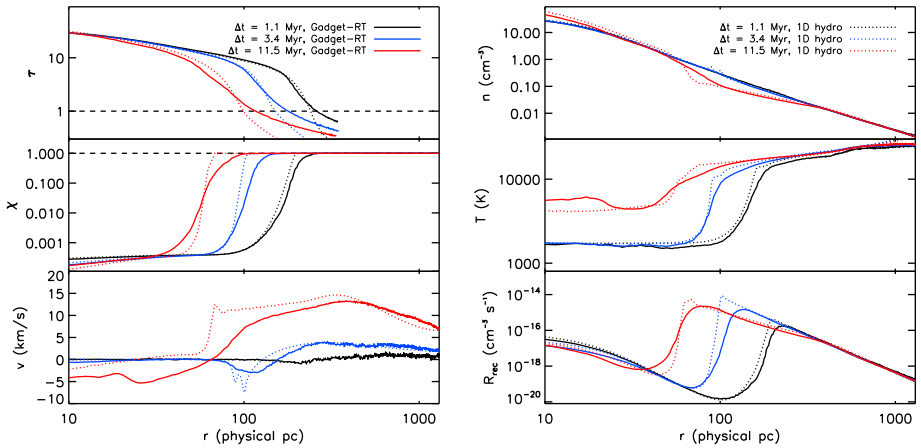

Figure 1.2: Radial profiles of the effective optical depth (top left panel), ionization fraction (middle left panel), radial velocity (bottom left panel), density (top right panel), gas temperature (middle left panel) and recombination rate (bottom right panel) from GADGET-RT of this work (solid) and the 1D radiation-hydro code of Ahn and Shapiro 2007 (dotted). The results are compared for $\Delta t = 1.1$ Myr (black), 3.4 Myr (blue) and 11.5 Myr (red). The radius on the x-axis is in the *physical* unit.

radial shells linearly spaced from $r = 10^{-3}r_t$ to $10r_t$. This is the same spatial resolution adopted in [2]. We bound the outer-most shell with the pressure of that shell that the initial condition, which becomes practically negligible as soon as ionization of outer shells photo-heats the gas.

We, then, create the corresponding initial condition for the GADGET-RT code. We set the box-size to be $20r_t$ and locate the center of the halo at the center of the box. We randomly place particles using the density profile as the probability function. The effective pressure for dark matter is converted into random velocity dispersion following the Boltzmann distribution.

[2] adopted radial EIBR that optical depth to the radiation at ith shell from the center at frequency ν is given by

$$\tau_{\nu,i} = dr \sum_{j=i}^{N_{\rm sh}} \sum_{X} n_X(r_j)\sigma_{X,\nu}, \tag{1.1}$$

where dr is the thickness of the shell, n_X is number density of species X, $\sigma_{X,\nu}$ is the cross-section of the species X for frequency ν, and the baryonic species X include HI, HeI, HeII, H^-, H2I and H2II. In order to reproduce isotropic EIBR that GADGET-RT code adopts, we modify the 1D code so that the optical depth is given by angular average for different lines of sights, $\vec{l}$:

$$\tau_{\nu,i} = (4\pi)^{-1} \int d\Omega \int_{l=0}^{l_{\rm max}} dl \sum_{X} n_{X,j}(r)\sigma_{X,\nu}. \tag{1.2}$$

Here, $l_{\rm max}$ is the distance from the ith shell to the outer-most shell. r is given by

$$r = \sqrt{r_i^2 + l^2 + 2lr_i\mu}, \tag{1.3}$$

where $\mu = \hat{l} \cdot \hat{r}$. Solving Equation (1.3) for l setting $r = r_{N_{\rm sh}}$ gives $l_{\rm max}$. See Figure ?? for the schematic description.

Since the angular integral in Equation (1.4) is symmetric for azimuthal direction, it can be simplified as following.

$$\tau_{\nu,i} = \frac{1}{2} \int_{-1}^{1} d\mu \int_{l=0}^{l_{\rm max}} dl \sum_{X} n_X(r) \sigma_{X,\nu}. \tag{1.4}$$

We calculate use interpolation to define $n_X(r)$ for $r_1 < r < r_{N_{\rm sh}}$. We use the Simpson's Rule to evaluate integrals. For the EIBR, we adopt the same parameters used in the standard case M_I0_z10 that the spectrum is given by 10^5 K blackbody spectrum and $J_{21} = 1$.

In Figure 1.1, we show particle maps with particle velocities and ionization status for four snapshots at Δt =0 (upper left panel), 3.43 (upper right panel), 5.73 (lower left panel) and 11.5 Myr (lower right panel) in the GADGET-RT simulation. The group of black arrows denote the central part of the halo that is shielded from the EIBR. The boundary between the region of black arrows and that of red arrows denote the ionization front that the gas is being ionized. The simulation clearly reproduces the propagation of the ionization front toward the halo center. We also find the ionization shock forming behind the front advancing toward the center over time shown as a ring of inward black arrows. At Δ = 11.5 Myr (right bottom panel), SPH particles outside of the halo clearly show an outflow. These features all agree qualitatively with findings in [2].

To make a quantitative comparison, we compare radial profiles of six

physical quantities in both simulations in Figure 1.2. The effective optical depth, τ, in the top right panel is defined by $\tau_{\rm eff} = -\log({\rm Tr})$ where ${\rm Tr} = (1/6)\Sigma_{\rm X=\pm x,y,z} \exp(-{\rm N_X}\sigma)$ is the average transmitivity from the six column densities for $\pm x, \pm y, \pm z$ directions calculated in the simulation. For the 1D code, the effective optical depth can be calculated precisely from the neutral hydrogen density profile. We also plot for ionized fraction (middle left panel), radial velocity (bottom left panel), density (top right panel), temperature (middle right panel), and recombination rate from Equation (2.2) (bottom right panel).

$\tau_{\rm eff}$ is somewhat overestimated in the outer part of the minihalo. This is because the 6-pixel sky approximation always results in a complete saturation of at least one sky-pixel toward the direction the minihalo while, in reality, the pixel should not be completely shielded as the angular size of the minihalo will be smaller that the sky-pixel and some ionizing background should be able to transmit from that direction. The overestimation of $\tau_{\rm eff}$, however, happens only at the region that the gas is already fully ionized and, therefore, has a negligible impact on the evaporation process. Other than that agreement between the two different code is reasonable.

For other quantities plotted in Figure 1.2, we, in general, find a good agreement between the two code. Transition of those quantities at the ionization front is more spread out in in the GADGET-RT code as the particle resolution in the GADGET-RT code is unable to perfectly reproduce sharp features in the 1D code. However, the outer part of the halo that is fully ion-

ized shows an excellent agreement in those quantities. The difference in the detailed features of I-front accumulates over time, but the agreement in the outer part remains good in later time. For the purpose of looking into the fate of ionized gas behind the ionization front that this work focuses on, the test result guarantees the reliability of the GADGET-RT code.

Appendix 2

Transverse Momentum Power Spectrum in Perturbation Theory

@*Transverse Momentum Power Spectrum in Perturbation Theory*

2.1 Low-k limit of the Ostriker-Vishniac Spectrum

Here, I derive the low-k limit of the OV spectrum (Eq. 3.28). First, I define $\epsilon \equiv k/k'$ and treat ϵ as a small number to find the leading order expression in ϵ. Then, the OV spectrum (Eq. 3.27) can be written as following.

$$\begin{aligned}
P_{q_\perp}^{\mathrm{OV,k\to 0}}(k,z) &= \dot{a}^2 f^2 \int \frac{k'^2 dk' d\mu'}{(2\pi)^2} P(|\mathbf{k}-\mathbf{k}'|,z) P(k',z) \frac{kk'(\epsilon-2\mu')(1-\mu'^2)}{k'^4(\epsilon^2+1-2\epsilon\mu')} \\
&= \dot{a}^2 f^2 \int \frac{dk' d\mu'}{(2\pi)^2} P(|\mathbf{k}-\mathbf{k}'|,z) P(k',z) \frac{k(\epsilon-2\mu')(1-\mu'^2)}{k'(\epsilon^2+1-2\epsilon\mu')} \\
&= \dot{a}^2 f^2 \int \frac{dk' d\mu'}{(2\pi)^2} (1-\mu'^2) P(|\mathbf{k}-\mathbf{k}'|,z) P(k',z) \epsilon(-2\mu'+\epsilon)(1+2\epsilon\mu') \\
&= \dot{a}^2 f^2 \int \frac{dk' d\mu'}{(2\pi)^2} (1-\mu'^2) P(|\mathbf{k}-\mathbf{k}'|,z) P(k',z) \epsilon\{-2\mu'+\epsilon(1-2\mu'^2)\} \qquad (2.1)
\end{aligned}$$

In the curly bracket of the above expression, $-2\mu'$ seems to be the leading order in ϵ. It, however, vanishes in the integration, $\int_{-1}^{1} d\mu'$, bringing up $\epsilon(1-2\mu'^2)$ as the true leading order.

Here, one should not neglect that $P(|\mathbf{k}-\mathbf{k}'|,z)$ also provides a term

that is first order in ϵ.

$$\begin{aligned} P(|\mathbf{k}-\mathbf{k}'|,z) &= P(k',z) - k\mu' \frac{dP}{dk'} \\ &= P(k',z)\left(1 - \epsilon\mu' \frac{d(\ln P)}{d(\ln k')}\right) \end{aligned} \tag{2.2}$$

Note that $\frac{d(\ln P)}{d(\ln k')}$ is 1 in the low-k limit and -3 in the high-k limit. We need to take into account above term as well to correctly derive the low-k limit of the OV spectrum.

$$\begin{aligned} P_{q_\perp}^{\mathrm{OV,k\to 0}}(k,z) &= \dot{a}^2 f^2 \int \frac{dk'd\mu'}{(2\pi)^2}(1-\mu'^2)\left[P(k',z) - k\mu'\frac{dP}{dk'}\right] P(k',z) \\ &\quad \epsilon\{-2\mu' + \epsilon(1-2\mu'^2)\} \\ &= \dot{a}^2 f^2 \int \frac{dk'd\mu'}{(2\pi)^2}(1-\mu'^2)P(k',z)^2\epsilon^2(1-2\mu'^2) \\ &+ \dot{a}^2 f^2 \int \frac{dk'd\mu'}{(2\pi)^2}(1-\mu'^2)P(k',z)\frac{dP}{dk'}2\epsilon k\mu'^2 \end{aligned} \tag{2.3}$$

In the above, I am only collecting terms in the second order in ϵ. The second part of integral in Equation (2.3) can transformed as following using integration by part.

$$\begin{aligned} P_{q_\perp}^{\mathrm{OV,k\to 0}}(k,z) &= \dot{a}^2 f^2 \int \frac{dk'd\mu'}{(2\pi)^2} P(k',z)^2\epsilon^2(1-2\mu'^2)(1-\mu'^2) \\ &+ \dot{a}^2 f^2 \int \frac{dk'd\mu'}{(2\pi)^2} P(k',z)^2\epsilon^2\mu'^2(1-\mu'^2) \end{aligned} \tag{2.4}$$

Integration w.r.t. μ' can be done analytically to get our final expression, Equation 3.28:

$$P_{q_\perp}^{\mathrm{OV,k\to 0}}(k,z) = \dot{a}^2 f^2 \left(\frac{8}{15}\right) \int \frac{dk'}{(2\pi)^2}\left(\frac{k^2}{k'^2}\right) P(k',z)^2.$$

2.2 Next-to-leading Order Connected term

We take the third term in Equation (3.23) and express it in terms of $\theta \equiv \nabla \cdot \mathbf{v}$, which is equivalent to say $\tilde{\mathbf{v}}(\mathbf{k}) = i(\hat{\mathbf{k}}/k)\tilde{\theta}$:

$$\begin{aligned} P_{q_\perp,c} &= \int \frac{d^3k'}{(2\pi)^3} \int \frac{d^3k''}{(2\pi)^3} [\hat{\mathbf{k}'} \cdot \hat{\mathbf{k}''} - (\hat{\mathbf{k}} \cdot \hat{\mathbf{k}'})(\hat{\mathbf{k}} \cdot \hat{\mathbf{k}''})] P_{\delta\delta vv,c}(\mathbf{k}-\mathbf{k}', -\mathbf{k}-\mathbf{k}'', \mathbf{k}', \mathbf{k}'') \\ &= -(\dot{a}f)^2 \int \frac{d^3k'}{(2\pi)^3} \int \frac{d^3k''}{(2\pi)^3} \left[\frac{\hat{\mathbf{k}'} \cdot \hat{\mathbf{k}''} - (\hat{\mathbf{k}} \cdot \hat{\mathbf{k}'})(\hat{\mathbf{k}} \cdot \hat{\mathbf{k}''})}{k'k''} \right] \\ &\quad P_{\delta\delta\theta\theta,c}(\mathbf{k}-\mathbf{k}', -\mathbf{k}-\mathbf{k}'', \mathbf{k}', \mathbf{k}''). \end{aligned} \tag{2.5}$$

Let us derive $P_{\delta\delta\theta\theta,c}$ using perturbation theory. We begin with the next-to-leading order expression for the full expression for $P_{\delta\delta\theta\theta}$ that includes the unconnected terms:

$$\begin{aligned} & (2\pi)^3 P_{\delta\delta\theta\theta}(\mathbf{k_1}, \mathbf{k_2}, \mathbf{k_3}, \mathbf{k_4}) \delta_{\rm D}(\mathbf{k_1} + \mathbf{k_2} + \mathbf{k_3} + \mathbf{k_4}) \\ = & \left\langle \tilde{\delta}(\mathbf{k_1}) \tilde{\delta}(\mathbf{k_2}) \tilde{\theta}(\mathbf{k_3}) \tilde{\theta}(\mathbf{k_4}) \right\rangle \\ = & \left\langle \tilde{\delta}^{(1)}(\mathbf{k_1}) \tilde{\delta}^{(2)}(\mathbf{k_2}) \tilde{\theta}^{(2)}(\mathbf{k_3}) \tilde{\theta}^{(1)}(\mathbf{k_4}) \right\rangle + \text{ cyclic (6 terms)} \\ + & \left\langle \tilde{\delta}^{(3)}(\mathbf{k_1}) \tilde{\delta}^{(1)}(\mathbf{k_2}) \tilde{\theta}^{(1)}(\mathbf{k_3}) \tilde{\theta}^{(1)}(\mathbf{k_4}) \right\rangle + \text{ cyclic (4 terms)}. \end{aligned} \tag{2.6}$$

The numbers in the superscripts indicate the perturbation order. We refer to the first case as P_{1221} term and the second case as P_{3111} term. Note that above expression *is not* symmetric for switching one of $\mathbf{k_1}$ and $\mathbf{k_2}$ with one of $\mathbf{k_3}$ and $\mathbf{k_4}$ although it *is* symmetric for $\mathbf{k_1} \leftrightarrow \mathbf{k_2}$ and $\mathbf{k_3} \leftrightarrow \mathbf{k_4}$.

One of P_{1221} terms, $\left\langle \tilde{\delta}^{(1)}(\mathbf{k_1})\tilde{\delta}^{(2)}(\mathbf{k_2})\tilde{\theta}^{(2)}(\mathbf{k_3})\tilde{\theta}^{(1)}(\mathbf{k_4})\right\rangle$, is given by

$$
\begin{aligned}
&\left\langle \tilde{\delta}^{(1)}(\mathbf{k_1})\tilde{\delta}^{(2)}(\mathbf{k_2})\tilde{\theta}^{(2)}(\mathbf{k_3})\tilde{\theta}^{(1)}(\mathbf{k_4})\right\rangle \\
= &\int \frac{d^3\mathbf{k_{2,a}}}{(2\pi)^3}\int d^3\mathbf{k_{2,b}}\int \frac{d^3\mathbf{k_{3,a}}}{(2\pi)^3}\int d^3\mathbf{k_{3,b}} \\
&F_2^{(s)}(\mathbf{k_{2,a}},\mathbf{k_{2,b}})G_2^{(s)}(\mathbf{k_{3,a}},\mathbf{k_{3,b}})\delta_{\rm D}(\mathbf{k_2}-\mathbf{k_{2,a}}-\mathbf{k_{2,b}})\delta_{\rm D}(\mathbf{k_3}-\mathbf{k_{3,a}}-\mathbf{k_{3,b}}) \\
&\left\langle \tilde{\delta}^{(1)}(\mathbf{k_1})\tilde{\delta}^{(1)}(\mathbf{k_{2,a}})\tilde{\delta}^{(1)}(\mathbf{k_{2,b}})\tilde{\delta}^{(1)}(\mathbf{k_{3,a}})\tilde{\delta}^{(1)}(\mathbf{k_{3,b}})\tilde{\delta}^{(1)}(\mathbf{k_4})\right\rangle \\
= &\int \frac{d^3\mathbf{k_{2,a}}}{(2\pi)^3}\int d^3\mathbf{k_{2,b}}\int \frac{d^3\mathbf{k_{3,a}}}{(2\pi)^3}\int d^3\mathbf{k_{3,b}} \\
&F_2^{(s)}(\mathbf{k_{2,a}},\mathbf{k_{2,b}})G_2^{(s)}(\mathbf{k_{3,a}},\mathbf{k_{3,b}})\delta_{\rm D}(\mathbf{k_2}-\mathbf{k_{2,a}}-\mathbf{k_{2,b}})\delta_{\rm D}(\mathbf{k_3}-\mathbf{k_{3,a}}-\mathbf{k_{3,b}}) \\
&\Big[\left\langle \tilde{\delta}^{(1)}(\mathbf{k_1})\tilde{\delta}^{(1)}(\mathbf{k_{2,a}})\right\rangle\left\langle \tilde{\delta}^{(1)}(\mathbf{k_{2,b}})\tilde{\delta}^{(1)}(\mathbf{k_{3,a}})\right\rangle\left\langle \tilde{\delta}^{(1)}(\mathbf{k_{3,b}})\tilde{\delta}^{(1)}(\mathbf{k_4})\right\rangle \\
+ &\left\langle \tilde{\delta}^{(1)}(\mathbf{k_1})\tilde{\delta}^{(1)}(\mathbf{k_{2,a}})\right\rangle\left\langle \tilde{\delta}^{(1)}(\mathbf{k_{2,b}})\tilde{\delta}^{(1)}(\mathbf{k_{3,b}})\right\rangle\left\langle \tilde{\delta}^{(1)}(\mathbf{k_{3,a}})\tilde{\delta}^{(1)}(\mathbf{k_4})\right\rangle \\
+ &\left\langle \tilde{\delta}^{(1)}(\mathbf{k_1})\tilde{\delta}^{(1)}(\mathbf{k_{2,a}})\right\rangle\left\langle \tilde{\delta}^{(1)}(\mathbf{k_{2,b}})\tilde{\delta}^{(1)}(\mathbf{k_4})\right\rangle\left\langle \tilde{\delta}^{(1)}(\mathbf{k_{3,a}})\tilde{\delta}^{(1)}(\mathbf{k_{3,b}})\right\rangle \quad \text{(vanish)} \\
+ &\left\langle \tilde{\delta}^{(1)}(\mathbf{k_1})\tilde{\delta}^{(1)}(\mathbf{k_{2,b}})\right\rangle \quad \text{(... are similar to the above 3 lines)} \\
+ &\left\langle \tilde{\delta}^{(1)}(\mathbf{k_1})\tilde{\delta}^{(1)}(\mathbf{k_{3,a}})\right\rangle\left\langle \tilde{\delta}^{(1)}(\mathbf{k_{2,a}})\tilde{\delta}^{(1)}(\mathbf{k_{2,b}})\right\rangle\left\langle \tilde{\delta}^{(1)}(\mathbf{k_{3,b}})\tilde{\delta}^{(1)}(\mathbf{k_4})\right\rangle \quad \text{(vanish)} \\
+ &\left\langle \tilde{\delta}^{(1)}(\mathbf{k_1})\tilde{\delta}^{(1)}(\mathbf{k_{3,a}})\right\rangle\left\langle \tilde{\delta}^{(1)}(\mathbf{k_{2,a}})\tilde{\delta}^{(1)}(\mathbf{k_{3,b}})\right\rangle\left\langle \tilde{\delta}^{(1)}(\mathbf{k_{2,b}})\tilde{\delta}^{(1)}(\mathbf{k_4})\right\rangle \\
+ &\left\langle \tilde{\delta}^{(1)}(\mathbf{k_1})\tilde{\delta}^{(1)}(\mathbf{k_{3,a}})\right\rangle\left\langle \tilde{\delta}^{(1)}(\mathbf{k_{2,a}})\tilde{\delta}^{(1)}(\mathbf{k_4})\right\rangle\left\langle \tilde{\delta}^{(1)}(\mathbf{k_{2,b}})\tilde{\delta}^{(1)}(\mathbf{k_{3,b}})\right\rangle \\
+ &\left\langle \tilde{\delta}^{(1)}(\mathbf{k_1})\tilde{\delta}^{(1)}(\mathbf{k_{3,b}})\right\rangle \quad \text{(... are similar to the above 3 lines)} \\
+ &\left\langle \tilde{\delta}^{(1)}(\mathbf{k_1})\tilde{\delta}^{(1)}(\mathbf{k_4})\right\rangle \quad \text{(... are the unconnected terms)}\Big], \qquad (2.7)
\end{aligned}
$$

where $F_2^{(s)}(\mathbf{q_1},\mathbf{q_2}) = \frac{5}{7} + \frac{1}{2}\frac{\mathbf{q_1}\cdot\mathbf{q_2}}{q_1 q_2} + \frac{2}{7}\left(\frac{\mathbf{q_1}\cdot\mathbf{q_2}}{q_1 q_2}\right)^2$ and $G_2^{(s)}(\mathbf{q_1},\mathbf{q_2}) = \frac{3}{7} + \frac{1}{2}\frac{\mathbf{q_1}\cdot\mathbf{q_2}}{q_1 q_2} + \frac{4}{7}\left(\frac{\mathbf{q_1}\cdot\mathbf{q_2}}{q_1 q_2}\right)^2$. The case that $\tilde{\delta}^{(1)}(\mathbf{k_1})$ is paired with $\tilde{\delta}^{(1)}(\mathbf{k_{2,a}})$ is equivalent to the case that it is paired with $\tilde{\delta}^{(1)}(\mathbf{k_{2,b}})$, the case that $\tilde{\delta}^{(1)}(\mathbf{k_{2,a}})$ or $\tilde{\theta}^{(1)}(\mathbf{k_{3,a}})$ is paired with $\tilde{\delta}^{(1)}(\mathbf{k_{2,b}})$ or $\tilde{\theta}^{(1)}(\mathbf{k_{3,b}})$, respectively, vanishes, and the case that

$\tilde{\delta}^{(1)}(\mathbf{k_1})$ is paired with $\tilde{\delta}^{(1)}(\mathbf{k_4})$ belongs to the unconnected moment. We proceed with the connected terms only:

$$
\begin{aligned}
& \left\langle \tilde{\delta}^{(1)}(\mathbf{k_1})\tilde{\delta}^{(2)}(\mathbf{k_2})\tilde{\theta}^{(2)}(\mathbf{k_3})\tilde{\theta}^{(1)}(\mathbf{k_4}) \right\rangle_c \\
= \;& 4\int \frac{d^3\mathbf{k_{2,a}}}{(2\pi)^3}\int d^3\mathbf{k_{2,b}}\int \frac{d^3\mathbf{k_{3,a}}}{(2\pi)^3}\int d^3\mathbf{k_{3,b}} \\
& F_2^{(s)}(\mathbf{k_{2,a}},\mathbf{k_{2,b}})G_2^{(s)}(\mathbf{k_{3,a}},\mathbf{k_{3,b}})\delta_{\rm D}(\mathbf{k_2}-\mathbf{k_{2,a}}-\mathbf{k_{2,b}})\delta_{\rm D}(\mathbf{k_3}-\mathbf{k_{3,a}}-\mathbf{k_{3,b}}) \\
& \Big[\left\langle \tilde{\delta}(\mathbf{k_1})\tilde{\delta}(\mathbf{k_{2,a}})\right\rangle \left\langle \tilde{\delta}(\mathbf{k_{2,b}})\tilde{\delta}(\mathbf{k_{3,a}})\right\rangle \left\langle \tilde{\delta}(\mathbf{k_{3,b}})\tilde{\delta}(\mathbf{k_4})\right\rangle \\
+ \;& \left\langle \tilde{\delta}(\mathbf{k_1})\tilde{\delta}(\mathbf{k_{3,a}})\right\rangle \left\langle \tilde{\delta}(\mathbf{k_{2,a}})\tilde{\delta}(\mathbf{k_{3,b}})\right\rangle \left\langle \tilde{\delta}(\mathbf{k_{2,b}})\tilde{\delta}(\mathbf{k_4})\right\rangle\Big] \\
\\
= \;& 4\int \frac{d^3\mathbf{k_{2,a}}}{(2\pi)^3}\int d^3\mathbf{k_{2,b}}\int \frac{d^3\mathbf{k_{3,a}}}{(2\pi)^3}\int d^3\mathbf{k_{3,b}} \\
& F_2^{(s)}(\mathbf{k_{2,a}},\mathbf{k_{2,b}})G_2^{(s)}(\mathbf{k_{3,a}},\mathbf{k_{3,b}})\delta_{\rm D}(\mathbf{k_2}-\mathbf{k_{2,a}}-\mathbf{k_{2,b}})\delta_{\rm D}(\mathbf{k_3}-\mathbf{k_{3,a}}-\mathbf{k_{3,b}}) \\
& [(2\pi)^3P(\mathbf{k_1})\delta_{\rm D}(\mathbf{k_1}+\mathbf{k_{2,a}})(2\pi)^3P(\mathbf{k_{2,b}})\delta_{\rm D}(\mathbf{k_{2,b}}+\mathbf{k_{3,a}})(2\pi)^3P(\mathbf{k_4})\delta_{\rm D}(\mathbf{k_{3,b}}+\mathbf{k_4}) \\
+ \;& (2\pi)^3P(\mathbf{k_1})\delta_{\rm D}(\mathbf{k_1}+\mathbf{k_{3,a}})(2\pi)^3P(\mathbf{k_{2,a}})\delta_{\rm D}(\mathbf{k_{3,b}}+\mathbf{k_{2,a}})(2\pi)^3P(\mathbf{k_4})\delta_{\rm D}(\mathbf{k_{2,b}}+\mathbf{k_4})] \\
\\
= \;& 4(2\pi)^3\delta_{\rm D}(\mathbf{k_1}+\mathbf{k_2}+\mathbf{k_3}+\mathbf{k_4}) \\
& [F_2^{(s)}(\mathbf{k_1}+\mathbf{k_2},-\mathbf{k_1})G_2^{(s)}(\mathbf{k_1}+\mathbf{k_2},\mathbf{k_4})P(\mathbf{k_1})P(\mathbf{k_1}+\mathbf{k_2})P(\mathbf{k_4}) \\
+ \;& G_2^{(s)}(\mathbf{k_1}+\mathbf{k_3},-\mathbf{k_1})F_2^{(s)}(\mathbf{k_1}+\mathbf{k_3},\mathbf{k_4})P(\mathbf{k_1})P(\mathbf{k_1}+\mathbf{k_3})P(\mathbf{k_4})]. \qquad (2.8)
\end{aligned}
$$

Similarly,

$$\begin{aligned}
&\left\langle \tilde{\delta}^{(2)}(\mathbf{k_1})\tilde{\delta}^{(1)}(\mathbf{k_2})\tilde{\theta}^{(2)}(\mathbf{k_3})\tilde{\theta}^{(1)}(\mathbf{k_4})\right\rangle_c \\
= \;& 4(2\pi)^3\delta_{\rm D}(\mathbf{k_1}+\mathbf{k_2}+\mathbf{k_3}+\mathbf{k_4}) \\
& [F_2^{(s)}(\mathbf{k_2}+\mathbf{k_1},-\mathbf{k_2})G_2^{(s)}(\mathbf{k_2}+\mathbf{k_1},\mathbf{k_4})P(\mathbf{k_2})P(\mathbf{k_2}+\mathbf{k_1})P(\mathbf{k_4}) \\
+ \;& G_2^{(s)}(\mathbf{k_2}+\mathbf{k_3},-\mathbf{k_2})F_2^{(s)}(\mathbf{k_2}+\mathbf{k_3},\mathbf{k_4})P(\mathbf{k_2})P(\mathbf{k_2}+\mathbf{k_3})P(\mathbf{k_4})], \quad (2.9)
\end{aligned}$$

$$\begin{aligned}
&\left\langle \tilde{\delta}^{(1)}(\mathbf{k_1})\tilde{\delta}^{(2)}(\mathbf{k_2})\tilde{\theta}^{(1)}(\mathbf{k_3})\tilde{\theta}^{(2)}(\mathbf{k_4})\right\rangle_c \\
= \;& 4(2\pi)^3\delta_{\rm D}(\mathbf{k_1}+\mathbf{k_2}+\mathbf{k_3}+\mathbf{k_4}) \\
& [F_2^{(s)}(\mathbf{k_1}+\mathbf{k_2},-\mathbf{k_1})G_2^{(s)}(\mathbf{k_1}+\mathbf{k_2},\mathbf{k_3})P(\mathbf{k_1})P(\mathbf{k_1}+\mathbf{k_2})P(\mathbf{k_3}) \\
+ \;& G_2^{(s)}(\mathbf{k_1}+\mathbf{k_4},-\mathbf{k_1})F_2^{(s)}(\mathbf{k_1}+\mathbf{k_4},\mathbf{k_3})P(\mathbf{k_1})P(\mathbf{k_1}+\mathbf{k_4})P(\mathbf{k_3})], \quad (2.10)
\end{aligned}$$

and

$$\begin{aligned}
&\left\langle \tilde{\delta}^{(2)}(\mathbf{k_1})\tilde{\delta}^{(1)}(\mathbf{k_2})\tilde{\theta}^{(1)}(\mathbf{k_3})\tilde{\theta}^{(2)}(\mathbf{k_4})\right\rangle_c \\
= \;& 4(2\pi)^3\delta_{\rm D}(\mathbf{k_1}+\mathbf{k_2}+\mathbf{k_3}+\mathbf{k_4}) \\
& [F_2^{(s)}(\mathbf{k_2}+\mathbf{k_1},-\mathbf{k_2})G_2^{(s)}(\mathbf{k_2}+\mathbf{k_1},\mathbf{k_3})P(\mathbf{k_2})P(\mathbf{k_2}+\mathbf{k_1})P(\mathbf{k_3}) \\
+ \;& G_2^{(s)}(\mathbf{k_2}+\mathbf{k_4},-\mathbf{k_2})F_2^{(s)}(\mathbf{k_2}+\mathbf{k_4},\mathbf{k_3})P(\mathbf{k_2})P(\mathbf{k_2}+\mathbf{k_4})P(\mathbf{k_3})]. \quad (2.11)
\end{aligned}$$

The following two terms can be expanded in the same way as well, but have slightly different forms:

$$\begin{aligned}
&\left\langle \tilde{\delta}^{(2)}(\mathbf{k_1})\tilde{\delta}^{(2)}(\mathbf{k_2})\tilde{\theta}^{(1)}(\mathbf{k_3})\tilde{\theta}^{(1)}(\mathbf{k_4})\right\rangle_c \\
= \;& 4(2\pi)^3\delta_{\rm D}(\mathbf{k_1}+\mathbf{k_2}+\mathbf{k_3}+\mathbf{k_4}) \\
& [F_2^{(s)}(\mathbf{k_3}+\mathbf{k_1},-\mathbf{k_3})F_2^{(s)}(\mathbf{k_3}+\mathbf{k_1},\mathbf{k_4})P(\mathbf{k_3})P(\mathbf{k_3}+\mathbf{k_1})P(\mathbf{k_4}) \\
+ \;& F_2^{(s)}(\mathbf{k_3}+\mathbf{k_2},-\mathbf{k_3})F_2^{(s)}(\mathbf{k_3}+\mathbf{k_2},\mathbf{k_4})P(\mathbf{k_3})P(\mathbf{k_2}+\mathbf{k_3})P(\mathbf{k_4})], \quad (2.12)
\end{aligned}$$

$$
\begin{aligned}
&\left\langle \tilde{\delta}^{(1)}(\mathbf{k_1})\tilde{\delta}^{(1)}(\mathbf{k_2})\tilde{\theta}^{(2)}(\mathbf{k_3})\tilde{\theta}^{(2)}(\mathbf{k_4})\right\rangle_c \\
= \;& 4(2\pi)^3\delta_{\rm D}(\mathbf{k_1}+\mathbf{k_2}+\mathbf{k_3}+\mathbf{k_4}) \\
& [G_2^{(s)}(\mathbf{k_1}+\mathbf{k_3},-\mathbf{k_1})G_2^{(s)}(\mathbf{k_1}+\mathbf{k_3},\mathbf{k_2})P(\mathbf{k_1})P(\mathbf{k_1}+\mathbf{k_3})P(\mathbf{k_2}) \\
+ \;& G_2^{(s)}(\mathbf{k_1}+\mathbf{k_4},-\mathbf{k_1})G_2^{(s)}(\mathbf{k_1}+\mathbf{k_4},\mathbf{k_2})P(\mathbf{k_1})P(\mathbf{k_1}+\mathbf{k_4})P(\mathbf{k_2})]. \qquad (2.13)
\end{aligned}
$$

For P_{3111} terms, we have

$$
\begin{aligned}
&\left\langle \tilde{\delta}^{(3)}(\mathbf{k_1})\tilde{\delta}^{(1)}(\mathbf{k_2})\tilde{\theta}^{(1)}(\mathbf{k_3})\tilde{\theta}^{(1)}(\mathbf{k_4})\right\rangle \\
= \;& \int\frac{dk_{1,a}^3}{(2\pi)^3}\int\frac{dk_{1,b}^3}{(2\pi)^3}\int dk_{1,c}^3\, F_3(\mathbf{k_{1,a}},\mathbf{k_{1,b}},\mathbf{k_{1,c}})\delta_D(\mathbf{k_1}-\mathbf{k_{1,a}}-\mathbf{k_{1,b}}-\mathbf{k_{1,c}}) \\
& \left\langle \tilde{\delta}^{(1)}(\mathbf{k_{1,a}})\tilde{\delta}^{(1)}(\mathbf{k_{1,b}})\tilde{\delta}^{(1)}(\mathbf{k_{1,c}})\tilde{\delta}^{(1)}(\mathbf{k_2})\tilde{\delta}^{(1)}(\mathbf{k_3})\tilde{\delta}^{(1)}(\mathbf{k_4})\right\rangle \\
= \;& \int\frac{dk_{1,a}^3}{(2\pi)^3}\int\frac{dk_{1,b}^3}{(2\pi)^3}\int dk_{1,c}^3\, F_3(\mathbf{k_{1,a}},\mathbf{k_{1,b}},\mathbf{k_{1,c}})\delta_D(\mathbf{k_1}-\mathbf{k_{1,a}}-\mathbf{k_{1,b}}-\mathbf{k_{1,c}}) \\
& \left\langle \tilde{\delta}^{(1)}(\mathbf{k_{1,a}})\tilde{\delta}^{(1)}(\mathbf{k_{1,b}})\tilde{\delta}^{(1)}(\mathbf{k_{1,c}})\tilde{\delta}^{(1)}(\mathbf{k_2})\tilde{\delta}^{(1)}(\mathbf{k_3})\tilde{\delta}^{(1)}(\mathbf{k_4})\right\rangle \\
= \;& 6(2\pi)^3F_3^{(s)}(-\mathbf{k_2},-\mathbf{k_3},-\mathbf{k_4})P(\mathbf{k_2})P(\mathbf{k_3})P(\mathbf{k_4})\delta_D(\mathbf{k_1}+\mathbf{k_2}+\mathbf{k_3}+\mathbf{k_4}), \qquad (2.14)
\end{aligned}
$$

$$
\begin{aligned}
&\left\langle \tilde{\delta}^{(1)}(\mathbf{k_1})\tilde{\delta}^{(3)}(\mathbf{k_2})\tilde{\theta}^{(1)}(\mathbf{k_3})\tilde{\theta}^{(1)}(\mathbf{k_4})\right\rangle \\
= \;& 6(2\pi)^3F_3^{(s)}(-\mathbf{k_1},-\mathbf{k_3},-\mathbf{k_4})P(\mathbf{k_1})P(\mathbf{k_3})P(\mathbf{k_4})\delta_D(\mathbf{k_1}+\mathbf{k_2}+\mathbf{k_3}+\mathbf{k_4}), \qquad (2.15)
\end{aligned}
$$

$$
\begin{aligned}
&\left\langle \tilde{\delta}^{(1)}(\mathbf{k_1})\tilde{\delta}^{(1)}(\mathbf{k_2})\tilde{\theta}^{(3)}(\mathbf{k_3})\tilde{\theta}^{(1)}(\mathbf{k_4})\right\rangle \\
= \;& 6(2\pi)^3G_3^{(s)}(-\mathbf{k_1},-\mathbf{k_2},-\mathbf{k_4})P(\mathbf{k_1})P(\mathbf{k_2})P(\mathbf{k_4})\delta_D(\mathbf{k_1}+\mathbf{k_2}+\mathbf{k_3}+\mathbf{k_4}), \qquad (2.16)
\end{aligned}
$$

and

$$
\begin{aligned}
&\left\langle \tilde{\delta}^{(1)}(\mathbf{k_1})\tilde{\delta}^{(1)}(\mathbf{k_2})\tilde{\theta}^{(1)}(\mathbf{k_3})\tilde{\theta}^{(3)}(\mathbf{k_4})\right\rangle \\
= \;& 6(2\pi)^3G_3^{(s)}(-\mathbf{k_1},-\mathbf{k_2},-\mathbf{k_3})P(\mathbf{k_1})P(\mathbf{k_2})P(\mathbf{k_3})\delta_D(\mathbf{k_1}+\mathbf{k_2}+\mathbf{k_3}+\mathbf{k_4}), \qquad (2.17)
\end{aligned}
$$

where the kernels, F_3 and G_3, are given by the following recursion relations.

$$\begin{aligned} F_n(\mathbf{q_1},...,\mathbf{q_n}) &= \sum_{m=1}^{n-1} \frac{G_m(\mathbf{q_1},...,\mathbf{q_m})}{(2n+3)(n-1)} \\ &\quad [(2n+1)\alpha(\mathbf{q_1}+...+\mathbf{q_m},\mathbf{q_{m+1}}+...+\mathbf{q_n})F_{n-m}(\mathbf{q_{m+1}},...,\mathbf{q_n}) \\ &\quad +2\beta(\mathbf{q_1}+...+\mathbf{q_m},\mathbf{q_{m+1}}+...+\mathbf{q_n})G_{n-m}(\mathbf{q_{m+1}},...,\mathbf{q_n})]\,, \end{aligned} \tag{2.18}$$

$$\begin{aligned} G_n(\mathbf{q_1},...,\mathbf{q_n}) &= \sum_{m=1}^{n-1} \frac{G_m(\mathbf{q_1},...,\mathbf{q_m})}{(2n+3)(n-1)} \\ &\quad [3\alpha(\mathbf{q_1}+...+\mathbf{q_m},\mathbf{q_{m+1}}+...+\mathbf{q_n})F_{n-m}(\mathbf{q_{m+1}},...,\mathbf{q_n}) \\ &\quad +2n\beta(\mathbf{q_1}+...+\mathbf{q_m},\mathbf{q_{m+1}}+...+\mathbf{q_n})G_{n-m}(\mathbf{q_{m+1}},...,\mathbf{q_n})]\,, \end{aligned} \tag{2.19}$$

and $F_3^{(s)}$ and $G_3^{(s)}$ denote the symmetrized kernels of F_3 and G_3.

Combining results above and substituting them into Equation (2.20), we obtain the expression for the transverse momentum power spectrum from

the connected terms:

$$
\begin{aligned}
P^{\rm NLO}_{q_\perp,c} &= -\int \frac{d^3k'}{(2\pi)^3}\int \frac{d^3k''}{(2\pi)^3}\left[\frac{\hat{\mathbf{k}'}\cdot\hat{\mathbf{k}''}-(\hat{\mathbf{k}}\cdot\hat{\mathbf{k}'})(\hat{\mathbf{k}}\cdot\hat{\mathbf{k}''})}{k'k''}\right] \\
&\quad \Big\{6\Big[F_3^{(s)}(\mathbf{k}+\mathbf{k}'',-\mathbf{k}',-\mathbf{k}'')P(\mathbf{k}+\mathbf{k}'')P(\mathbf{k}')P(\mathbf{k}'') \\
&+ F_3^{(s)}(\mathbf{k}-\mathbf{k}',\mathbf{k}',\mathbf{k}'')P(\mathbf{k}-\mathbf{k}')P(\mathbf{k}')P(\mathbf{k}'') \\
&+ G_3^{(s)}(\mathbf{k}-\mathbf{k}',-\mathbf{k}-\mathbf{k}'',\mathbf{k}'')P(\mathbf{k}-\mathbf{k}')P(\mathbf{k}+\mathbf{k}'')P(\mathbf{k}'') \\
&+ G_3^{(s)}(\mathbf{k}-\mathbf{k}',-\mathbf{k}-\mathbf{k}'',\mathbf{k}')P(\mathbf{k}-\mathbf{k}')P(\mathbf{k}+\mathbf{k}'')P(\mathbf{k}')] \\
&+ 4[F_2^{(s)}(\mathbf{k}-\mathbf{k}',\mathbf{k}'+\mathbf{k}'')G_2^{(s)}(\mathbf{k}'+\mathbf{k}'',-\mathbf{k}'')P(\mathbf{k}-\mathbf{k}')P(\mathbf{k}'+\mathbf{k}'')P(\mathbf{k}'') \\
&+ F_2^{(s)}(\mathbf{k}-\mathbf{k}',\mathbf{k}'+\mathbf{k}'')G_2^{(s)}(-\mathbf{k}'-\mathbf{k}'',\mathbf{k}')P(\mathbf{k}-\mathbf{k}')P(\mathbf{k}'+\mathbf{k}'')P(\mathbf{k}') \\
&+ G_2^{(s)}(\mathbf{k},-\mathbf{k}+\mathbf{k}')F_2^{(s)}(\mathbf{k},\mathbf{k}'')P(\mathbf{k}-\mathbf{k}')P(\mathbf{k})P(\mathbf{k}'') \\
&+ G_2^{(s)}(\mathbf{k},-\mathbf{k}+\mathbf{k}')G_2^{(s)}(\mathbf{k},-\mathbf{k}-\mathbf{k}'')P(\mathbf{k}-\mathbf{k}')P(\mathbf{k})P(\mathbf{k}+\mathbf{k}'') \\
&+ G_2^{(s)}(\mathbf{k}-\mathbf{k}'+\mathbf{k}'',-\mathbf{k}+\mathbf{k}')F_2^{(s)}(\mathbf{k}-\mathbf{k}'+\mathbf{k}'',\mathbf{k}')P(\mathbf{k}-\mathbf{k}')P(\mathbf{k}-\mathbf{k}'+\mathbf{k}'')P(\mathbf{k}') \\
&+ G_2^{(s)}(\mathbf{k}-\mathbf{k}'+\mathbf{k}'',-\mathbf{k}+\mathbf{k}')G_2^{(s)}(\mathbf{k}-\mathbf{k}'+\mathbf{k}'',-\mathbf{k}-\mathbf{k}'') \\
&\quad P(\mathbf{k}-\mathbf{k}')P(\mathbf{k}-\mathbf{k}'+\mathbf{k}'')P(\mathbf{k}+\mathbf{k}'') \\
&+ F_2^{(s)}(-\mathbf{k}'-\mathbf{k}'',\mathbf{k}+\mathbf{k}'')G_2^{(s)}(-\mathbf{k}'-\mathbf{k}'',\mathbf{k}'')P(\mathbf{k}+\mathbf{k}'')P(\mathbf{k}'+\mathbf{k}'')P(\mathbf{k}'') \\
&+ F_2^{(s)}(-\mathbf{k}'-\mathbf{k}'',\mathbf{k}+\mathbf{k}'')G_2^{(s)}(-\mathbf{k}'-\mathbf{k}'',\mathbf{k}')P(\mathbf{k}+\mathbf{k}'')P(\mathbf{k}'+\mathbf{k}'')P(\mathbf{k}') \\
&+ G_2^{(s)}(-\mathbf{k}-\mathbf{k}''+\mathbf{k}',\mathbf{k}+\mathbf{k}'')F_2^{(s)}(-\mathbf{k}-\mathbf{k}''+\mathbf{k}',\mathbf{k}'')P(\mathbf{k}+\mathbf{k}'')P(\mathbf{k}-\mathbf{k}'+\mathbf{k}'')P(\mathbf{k}'') \\
&+ G_2^{(s)}(-\mathbf{k},\mathbf{k}+\mathbf{k}'')F_2^{(s)}(-\mathbf{k},\mathbf{k}')P(\mathbf{k}+\mathbf{k}'')P(\mathbf{k})P(\mathbf{k}') \\
&+ F_2^{(s)}(\mathbf{k},-\mathbf{k}')F_2^{(s)}(\mathbf{k},\mathbf{k}'')P(\mathbf{k})P(\mathbf{k}')P(\mathbf{k}'') \\
&+ F_2^{(s)}(\mathbf{k}-\mathbf{k}'+\mathbf{k}'',\mathbf{k}')F_2^{(s)}(\mathbf{k}-\mathbf{k}'+\mathbf{k}'',-\mathbf{k}'')P(\mathbf{k}')P(\mathbf{k}-\mathbf{k}'+\mathbf{k}'')P(\mathbf{k}'')\Big]\Big\} \quad (2.20)
\end{aligned}
$$

2.2.1 Low-k limit

Here, we derive the low-k limit of the connected term (Eq. 2.20). Figure 3.3 shows the connected term converging to a constant value in as $k \to 0$.

We start by setting $k = 0$ for the integrand of Eq. 2.20,

$$P_{\delta\delta\theta\theta}(-\mathbf{k}', -\mathbf{k}'', \mathbf{k}', \mathbf{k}'')\frac{\hat{\mathbf{k}'} \cdot \hat{\mathbf{k}''} - (\hat{\mathbf{k}} \cdot \hat{\mathbf{k}'})(\hat{\mathbf{k}} \cdot \hat{\mathbf{k}''})}{k'k''}. \tag{2.21}$$

Since the factor, $\frac{\hat{\mathbf{k}'} \cdot \hat{\mathbf{k}''} - (\hat{\mathbf{k}} \cdot \hat{\mathbf{k}'})(\hat{\mathbf{k}} \cdot \hat{\mathbf{k}''})}{k'k''}$, is anti-symmetric for $\mathbf{k}'$ and $\mathbf{k}''$. Thus, if a term that constitutes $P_{\delta\delta\theta\theta}$ is symmetric for $\mathbf{k}'$ or $\mathbf{k}''$, if will cancel in the

integral w.r.t. $\mathbf{k}'$ or $\mathbf{k}''$ and can be ignored in the derivation.

$$
\begin{aligned}
& \left[\frac{\hat{\mathbf{k}'}\cdot\hat{\mathbf{k}''}-(\hat{\mathbf{k}}\cdot\hat{\mathbf{k}'})(\hat{\mathbf{k}}\cdot\hat{\mathbf{k}''})}{k'k''}\right] P_{\delta\delta\theta\theta}(-\mathbf{k}',-\mathbf{k}'',\mathbf{k}',\mathbf{k}'') \\
= & \left[\frac{\hat{\mathbf{k}'}\cdot\hat{\mathbf{k}''}-(\hat{\mathbf{k}}\cdot\hat{\mathbf{k}'})(\hat{\mathbf{k}}\cdot\hat{\mathbf{k}''})}{k'k''}\right] \\
& \{6[F_3^{(s)}(-\mathbf{k}'',\mathbf{k}',\mathbf{k}'')P(\mathbf{k}')P(\mathbf{k}'')^2 \quad \text{(vanishes due to symmetry.)} \\
+ & F_3^{(s)}(-\mathbf{k}',\mathbf{k}',\mathbf{k}'')P(\mathbf{k}')^2P(\mathbf{k}'') \quad \text{(vanishes due to symmetry.)} \\
+ & G_3^{(s)}(-\mathbf{k}',-\mathbf{k}'',\mathbf{k}'')P(\mathbf{k}')P(\mathbf{k}'')^2) \quad \text{(vanishes due to symmetry.)} \\
+ & G_3^{(s)}(-\mathbf{k}',-\mathbf{k}'',\mathbf{k}')P(\mathbf{k}')^2P(\mathbf{k}'')] \quad \text{(vanishes due to symmetry.)} \\
+ & 4[F_2^{(s)}(\mathbf{k}'+\mathbf{k}'',-\mathbf{k}')G_2^{(s)}(\mathbf{k}'+\mathbf{k}'',-\mathbf{k}'')P(\mathbf{k}')P(\mathbf{k}'+\mathbf{k}'')P(\mathbf{k}'') \\
+ & F_2^{(s)}(\mathbf{k}'+\mathbf{k}'',-\mathbf{k}'')G_2^{(s)}(\mathbf{k}'+\mathbf{k}'',-\mathbf{k}')P(\mathbf{k}'')P(\mathbf{k}'+\mathbf{k}'')P(\mathbf{k}') \\
+ & F_2^{(s)}(\mathbf{k}'+\mathbf{k}'',-\mathbf{k}')G_2^{(s)}(\mathbf{k}'+\mathbf{k}'',-\mathbf{k}')P(\mathbf{k}')P(\mathbf{k}'+\mathbf{k}'')P(\mathbf{k}') \\
+ & F_2^{(s)}(\mathbf{k}'+\mathbf{k}'',-\mathbf{k}'')G_2^{(s)}(\mathbf{k}'+\mathbf{k}'',-\mathbf{k}'')P(\mathbf{k}'')P(\mathbf{k}'+\mathbf{k}'')P(\mathbf{k}'') \\
+ & F_2^{(s)}(\mathbf{k}'-\mathbf{k}'',-\mathbf{k}')G_2^{(s)}(\mathbf{k}'-\mathbf{k}'',-\mathbf{k}')P(\mathbf{k}')P(\mathbf{k}'-\mathbf{k}'')P(\mathbf{k}') \\
+ & F_2^{(s)}(\mathbf{k}'-\mathbf{k}'',\mathbf{k}'')G_2^{(s)}(\mathbf{k}'-\mathbf{k}'',\mathbf{k}'')P(\mathbf{k}'')P(\mathbf{k}'-\mathbf{k}'')P(\mathbf{k}'') \\
+ & G_2^{(s)}(\mathbf{k}'-\mathbf{k}'',-\mathbf{k}')G_2^{(s)}(\mathbf{k}'-\mathbf{k}'',\mathbf{k}'')P(\mathbf{k}')P(\mathbf{k}'-\mathbf{k}'')P(\mathbf{k}'') \\
+ & F_2^{(s)}(\mathbf{k}'-\mathbf{k}'',-\mathbf{k}')F_2^{(s)}(\mathbf{k}'-\mathbf{k}'',\mathbf{k}'')P(\mathbf{k}')P(\mathbf{k}'-\mathbf{k}'')P(\mathbf{k}'')]\}
\end{aligned}
\tag{2.22}
$$

Now, we make $\mathbf{k}'' \to -\mathbf{k}''$ transformation to group terms in a convenient way. Note that due to the anti-symmetric factor, $\left[\frac{\hat{\mathbf{k}'}\cdot\hat{\mathbf{k}''}-(\hat{\mathbf{k}}\cdot\hat{\mathbf{k}'})(\hat{\mathbf{k}}\cdot\hat{\mathbf{k}''})}{k'k''}\right]$, we should

flip the sign for each transformation. We omit the factor below for simplicity.

$$
\begin{aligned}
& P_{\delta\delta\theta\theta}(-\mathbf{k}', -\mathbf{k}'', \mathbf{k}', \mathbf{k}'') \\
= \; & 4[F_2^{(s)}(\mathbf{k}'+\mathbf{k}'', -\mathbf{k}')G_2^{(s)}(\mathbf{k}'+\mathbf{k}'', -\mathbf{k}'')P(\mathbf{k}')P(\mathbf{k}'+\mathbf{k}'')P(\mathbf{k}'') \\
+ \; & F_2^{(s)}(\mathbf{k}'+\mathbf{k}'', -\mathbf{k}'')G_2^{(s)}(\mathbf{k}'+\mathbf{k}'', -\mathbf{k}')P(\mathbf{k}'')P(\mathbf{k}'+\mathbf{k}'')P(\mathbf{k}') \\
+ \; & F_2^{(s)}(\mathbf{k}'+\mathbf{k}'', -\mathbf{k}')G_2^{(s)}(\mathbf{k}'+\mathbf{k}'', -\mathbf{k}')P(\mathbf{k}')P(\mathbf{k}'+\mathbf{k}'')P(\mathbf{k}') \\
+ \; & F_2^{(s)}(\mathbf{k}'+\mathbf{k}'', -\mathbf{k}'')G_2^{(s)}(\mathbf{k}'+\mathbf{k}'', -\mathbf{k}'')P(\mathbf{k}'')P(\mathbf{k}'+\mathbf{k}'')P(\mathbf{k}'') \\
- \; & F_2^{(s)}(\mathbf{k}'+\mathbf{k}'', -\mathbf{k}')G_2^{(s)}(\mathbf{k}'+\mathbf{k}'', -\mathbf{k}')P(\mathbf{k}')P(\mathbf{k}'+\mathbf{k}'')P(\mathbf{k}') \\
- \; & F_2^{(s)}(\mathbf{k}'+\mathbf{k}'', -\mathbf{k}'')G_2^{(s)}(\mathbf{k}'+\mathbf{k}'', -\mathbf{k}'')P(\mathbf{k}'')P(\mathbf{k}'+\mathbf{k}'')P(\mathbf{k}'') \\
- \; & G_2^{(s)}(\mathbf{k}'+\mathbf{k}'', -\mathbf{k}')G_2^{(s)}(\mathbf{k}'+\mathbf{k}'', -\mathbf{k}'')P(\mathbf{k}')P(\mathbf{k}'+\mathbf{k}'')P(\mathbf{k}'') \\
- \; & F_2^{(s)}(\mathbf{k}'+\mathbf{k}'', -\mathbf{k}')F_2^{(s)}(\mathbf{k}'+\mathbf{k}'', -\mathbf{k}'')P(\mathbf{k}')P(\mathbf{k}'+\mathbf{k}'')P(\mathbf{k}'')
\end{aligned}
$$

$$
\begin{aligned}
= \; & 4P(\mathbf{k}')P(\mathbf{k}'+\mathbf{k}'')P(\mathbf{k}'') \\
& [F_2^{(s)}(\mathbf{k}'+\mathbf{k}'', -\mathbf{k}'')G_2^{(s)}(\mathbf{k}'+\mathbf{k}'', -\mathbf{k}') + F_2^{(s)}(\mathbf{k}'+\mathbf{k}'', -\mathbf{k}'')G_2^{(s)}(\mathbf{k}'+\mathbf{k}'', -\mathbf{k}') \\
& -G_2^{(s)}(\mathbf{k}'+\mathbf{k}'', -\mathbf{k}')G_2^{(s)}(\mathbf{k}'+\mathbf{k}'', -\mathbf{k}'') - F_2^{(s)}(\mathbf{k}'+\mathbf{k}'', -\mathbf{k}')F_2^{(s)}(\mathbf{k}'+\mathbf{k}'', -\mathbf{k}'')]
\end{aligned}
$$

$$
\begin{aligned}
= \; & 4P(\mathbf{k}')P(\mathbf{k}'+\mathbf{k}'')P(\mathbf{k}'') \\
& [F_2^{(s)}(\mathbf{k}'+\mathbf{k}'', -\mathbf{k}'')[G_2^{(s)}(\mathbf{k}'+\mathbf{k}'', -\mathbf{k}') - F_2^{(s)}(\mathbf{k}'+\mathbf{k}'', -\mathbf{k}')] \\
+ \; & G_2^{(s)}(\mathbf{k}'+\mathbf{k}'', -\mathbf{k}')[F_2^{(s)}(\mathbf{k}'+\mathbf{k}'', -\mathbf{k}'') - G_2^{(s)}(\mathbf{k}'+\mathbf{k}'', -\mathbf{k}'')]]
\end{aligned}
$$

$$
\begin{aligned}
&= -4P(\mathbf{k}')P(\mathbf{k}'+\mathbf{k}'')P(\mathbf{k}'') \\
&\quad [F_2^{(s)}(\mathbf{k}'+\mathbf{k}'',-\mathbf{k}') - G_2^{(s)}(\mathbf{k}'+\mathbf{k}'',-\mathbf{k}')] \\
&\quad [F_2^{(s)}(\mathbf{k}'+\mathbf{k}'',-\mathbf{k}'') - G_2^{(s)}(\mathbf{k}'+\mathbf{k}'',-\mathbf{k}'')] \\
&= -4P(\mathbf{k}')P(\mathbf{k}'+\mathbf{k}'')P(\mathbf{k}'') \\
&\quad \left(\frac{2}{7}\right)^2 \left(1 - \frac{(\mathbf{k}'\cdot(\mathbf{k}'+\mathbf{k}''))^2}{k'^2(\mathbf{k}'+\mathbf{k}'')^2}\right)\left(1 - \frac{(\mathbf{k}''\cdot(\mathbf{k}'+\mathbf{k}''))^2}{k''^2(\mathbf{k}'+\mathbf{k}'')^2}\right) \\
&= -4\left(\frac{2}{7}\right)^2 P(\mathbf{k}')P(\mathbf{k}'+\mathbf{k}'')P(\mathbf{k}'') \\
&\quad \frac{k'^2(\mathbf{k}'+\mathbf{k}'')^2 - (k'^2+\mu k'k'')^2}{k'^2(\mathbf{k}'+\mathbf{k}'')^2}\left(1 - \frac{(\mathbf{k}''\cdot(\mathbf{k}'+\mathbf{k}''))^2}{k''^2(\mathbf{k}'+\mathbf{k}'')^2}\right) \\
&= -4\left(\frac{2}{7}\right)^2 P(\mathbf{k}')P(\mathbf{k}'+\mathbf{k}'')P(\mathbf{k}'') \\
&\quad \frac{(k'^4+2\mu k'^3k''+k'^2k''^2) - (k'^4+2\mu k'^3k''+\mu^2k'^2k''^2)}{k'^2(\mathbf{k}'+\mathbf{k}'')^2}\left(1 - \frac{(\mathbf{k}''\cdot(\mathbf{k}'+\mathbf{k}''))^2}{k''^2(\mathbf{k}'+\mathbf{k}'')^2}\right) \\
&= -4\left(\frac{2}{7}\right)^2 P(\mathbf{k}')P(\mathbf{k}'+\mathbf{k}'')P(\mathbf{k}'') \\
&\quad \frac{(1-\mu^2)k''^2}{(\mathbf{k}'+\mathbf{k}'')^2}\left(1 - \frac{(\mathbf{k}''\cdot(\mathbf{k}'+\mathbf{k}''))^2}{k''^2(\mathbf{k}'+\mathbf{k}'')^2}\right) \\
&= -4\left(\frac{2}{7}\right)^2 P(\mathbf{k}')P(\mathbf{k}'+\mathbf{k}'')P(\mathbf{k}'') \\
&\quad \left(\frac{(1-\mu^2)k''^2}{(\mathbf{k}'+\mathbf{k}'')^2}\right)\left(\frac{(1-\mu^2)k'^2}{(\mathbf{k}'+\mathbf{k}'')^2}\right) \qquad (2.23)
\end{aligned}
$$

$$= -4\left(\frac{2}{7}\right)^2 P(\mathbf{k}')P(\mathbf{k}'+\mathbf{k}'')P(\mathbf{k}'')\left(\frac{(1-\mu^2)k''k'}{(\mathbf{k}'+\mathbf{k}'')^2}\right)^2 \qquad (2.24)$$

Therefore, the low-k limit of $P_{q_\perp,c}$ becomes

$$\begin{aligned}
P_{q_\perp,c} &= -(\dot{a}f)^2\int\frac{d^3k'}{(2\pi)^3}\int\frac{d^3k''}{(2\pi)^3}\frac{\hat{\mathbf{k}}'\cdot\hat{\mathbf{k}}''-(\hat{\mathbf{k}}\cdot\hat{\mathbf{k}}')(\hat{\mathbf{k}}\cdot\hat{\mathbf{k}}'')}{k'k''}P_{\delta\delta\theta\theta}(-\mathbf{k}',-\mathbf{k}'',\mathbf{k}',\mathbf{k}'') \\
&= -(\dot{a}f)^2\int\frac{d^3k'}{(2\pi)^3}\int\frac{d^3k''}{(2\pi)^3}\frac{k'_xk''_x+k'_yk''_y}{k'^2k''^2}P_{\delta\delta\theta\theta}(-\mathbf{k}',-\mathbf{k}'',\mathbf{k}',\mathbf{k}'') \\
&= 4\left(\frac{2}{7}\right)^2(\dot{a}f)^2\int\frac{d^3k'}{(2\pi)^3}\int\frac{d^3k''}{(2\pi)^3}\frac{k'_xk''_x+k'_yk''_y}{k'^2k''^2}P(\mathbf{k}')P(\mathbf{k}'+\mathbf{k}'')P(\mathbf{k}'')\left(\frac{(1-\mu^2)k''k'}{(\mathbf{k}'+\mathbf{k}'')^2}\right)^2 \\
&= 4\left(\frac{2}{7}\right)^2(\dot{a}f)^2\int\frac{d^3k'}{(2\pi)^3}\int\frac{d^3k''}{(2\pi)^3}(k'_xk''_x+k'_yk''_y)P(\mathbf{k}')P(\mathbf{k}'+\mathbf{k}'')P(\mathbf{k}'')\frac{(1-\mu^2)^2}{(\mathbf{k}'+\mathbf{k}'')^4}, \qquad (2.25)
\end{aligned}$$

where we assume that the $\mathbf{k}$ vector is pointing at the $+z$-direction.

2.3 Next-to-leading Order Unconnected term

For the unconnected term, we begin by substituting the next-to-leading order power spectra of $P^{(2)}_{\delta\delta}$, $P^{(2)}_{\delta\theta}$ and $P^{(2)}_{\theta\theta}$ in Equation (3.26). Then, pertur-

bation theory gives

$$
\begin{aligned}
P^{(2)}_{\delta\delta}(\mathbf{k}) &= \int \frac{d^3k'}{(2\pi)^3} F_2^{(s)}(\mathbf{k}-\mathbf{k}',\mathbf{k}')^2 P(|\mathbf{k}-\mathbf{k}'|)P(k') \\
&+ 2\int \frac{d^3k'}{(2\pi)^3} F_3^{(s)}(\mathbf{k},-\mathbf{k}',\mathbf{k}')P(k)P(k'), \\
P^{(2)}_{\delta\theta}(\mathbf{k}) &= \int \frac{d^3k'}{(2\pi)^3} F_2^{(s)}(\mathbf{k}-\mathbf{k}',\mathbf{k}')G_2^{(s)}(\mathbf{k}-\mathbf{k}',\mathbf{k}')P(|\mathbf{k}-\mathbf{k}'|)P(k') \\
&+ \int \frac{d^3k'}{(2\pi)^3} F_3^{(s)}(\mathbf{k},-\mathbf{k}',\mathbf{k}')P(k)P(k') \\
&+ \int \frac{d^3k'}{(2\pi)^3} G_3^{(s)}(\mathbf{k},-\mathbf{k}',\mathbf{k}')P(k)P(k'), \\
P^{(2)}_{\theta\theta}(\mathbf{k}) &= \int \frac{d^3k'}{(2\pi)^3} G_2^{(s)}(\mathbf{k}-\mathbf{k}',\mathbf{k}')^2 P(|\mathbf{k}-\mathbf{k}'|)P(k') \\
&+ 2\int \frac{d^3k'}{(2\pi)^3} G_3^{(s)}(\mathbf{k},-\mathbf{k}',\mathbf{k}')P(k)P(k'). \qquad (2.26)
\end{aligned}
$$

Substituting $P_{\delta\delta} = P^{(2)}_{\delta\delta}$, $P_{\delta\theta} = P^{(2)}_{\delta\theta}$ and $P_{\theta\theta} = P^{(2)}_{\theta\theta}$ to Equation (3.26) and dropping higher order terms like $P^{(2)}P^{(2)}$, we obtain the expression for the transverse momentum power spectrum from the unconnected terms:

$$
\begin{aligned}
&P^{\rm NLO}_{q_\perp,uc}(k,z) = \dot{a}^2 f^2 \int \frac{d^3k'}{(2\pi)^3}(1-\mu'^2) \\
&\left\{ \frac{1}{k'^2}\left[P(|\mathbf{k}-\mathbf{k}'|)P^{(2)}_{\theta\theta}(k') + P^{(2)}_{\delta\delta}(|\mathbf{k}-\mathbf{k}'|)P(k') \right] + \right. \\
&\left. \frac{1}{|\mathbf{k}-\mathbf{k}'|^2}\left[P(|\mathbf{k}-\mathbf{k}'|)P^{(2)}_{\delta\theta}(k') + P^{(2)}_{\delta\theta}(|\mathbf{k}-\mathbf{k}'|)P(k') \right] \right\}. \qquad (2.27)
\end{aligned}
$$

Bibliography

[1] K. Ahn, I. T. Iliev, P. R. Shapiro, G. Mellema, J. Koda, and Y. Mao. Detecting the Rise and Fall of the First Stars by Their Impact on Cosmic Reionization. *ApJL*, 756:L16, September 2012.

[2] K. Ahn and P. R. Shapiro. Does radiative feedback by the first stars promote or prevent second generation star formation? *MNRAS*, 375:881–908, March 2007.

[3] R. Barkana and A. Loeb. Unusually Large Fluctuations in the Statistics of Galaxy Formation at High Redshift. *ApJ*, 609:474–481, July 2004.

[4] N. Battaglia, A. Natarajan, H. Trac, R. Cen, and A. Loeb. Reionization on Large Scales. III. Predictions for Low-l Cosmic Microwave Background Polarization and High-l Kinetic Sunyaev-Zel'dovich Observables. *ApJ*, 776:83, October 2013.

[5] N. Battaglia, H. Trac, R. Cen, and A. Loeb. Reionization on Large Scales. I. A Parametric Model Constructed from Radiation-hydrodynamic Simulations. *ApJ*, 776:81, October 2013.

[6] A. Bauer, V. Springel, M. Vogelsberger, S. Genel, P. Torrey, D. Sijacki, D. Nelson, and L. Hernquist. Hydrogen Reionization in the Illustris Universe. *ArXiv e-prints*, 1503.00734, March 2015.

[7] F. Bernardeau, S. Colombi, E. Gaztañaga, and R. Scoccimarro. Large-scale structure of the Universe and cosmological perturbation theory. *Phys. Rev. D*, 367:1–248, September 2002.

[8] R. J. Bouwens, L. Bradley, A. Zitrin, D. Coe, M. Franx, W. Zheng, R. Smit, O. Host, M. Postman, L. Moustakas, I. Labbé, M. Carrasco, A. Molino, M. Donahue, D. D. Kelson, M. Meneghetti, N. Benítez, D. Lemze, K. Umetsu, T. Broadhurst, J. Moustakas, P. Rosati, S. Jouvel, M. Bartelmann, H. Ford, G. Graves, C. Grillo, L. Infante, Y. Jimenez-Teja, O. Lahav, D. Maoz, E. Medezinski, P. Melchior, J. Merten, M. Nonino, S. Ogaz, and S. Seitz. A Census of Star-forming Galaxies in the z ˜ 9-10 Universe based on HST+Spitzer Observations over 19 Clash Clusters: Three Candidate z ˜ 9-10 Galaxies and Improved Constraints on the Star Formation Rate Density at z ˜ 9.2. *ApJ*, 795:126, November 2014.

[9] R. Cen and J. P. Ostriker. Where Are the Baryons? *ApJ*, 514:1–6, March 1999.

[10] B. Ciardi, J. S. Bolton, A. Maselli, and L. Graziani. The effect of inter-galactic helium on hydrogen reionization: implications for the sources of ionizing photons at z>6. *MNRAS*, 423:558–574, June 2012.

[11] S. Dodelson and J. M. Jubas. Reionization and its imprint of the cosmic microwave background. *ApJ*, 439:503–516, February 1995.

[12] J. D. Emberson, R. M. Thomas, and M. A. Alvarez. The Opacity of the

Intergalactic Medium during Reionization: Resolving Small-scale Structure. *ApJ*, 763:146, February 2013.

[13] X. Fan, M. A. Strauss, R. H. Becker, R. L. White, J. E. Gunn, G. R. Knapp, G. T. Richards, D. P. Schneider, J. Brinkmann, and M. Fukugita. Constraining the Evolution of the Ionizing Background and the Epoch of Reionization with z˜6 Quasars. II. A Sample of 19 Quasars. *AJ*, 132:117–136, July 2006.

[14] S. L. Finkelstein, R. E. Ryan, Jr., C. Papovich, M. Dickinson, M. Song, R. Somerville, H. C. Ferguson, B. Salmon, M. Giavalisco, A. M. Koekemoer, M. L. N. Ashby, P. Behroozi, M. Castellano, J. S. Dunlop, S. M. Faber, G. G. Fazio, A. Fontana, N. A. Grogin, N. Hathi, J. Jaacks, D. D. Kocevski, R. Livermore, R. J. McLure, E. Merlin, B. Mobasher, J. A. Newman, M. Rafelski, V. Tilvi, and S. P. Willner. The Evolution of the Galaxy Rest-Frame Ultraviolet Luminosity Function Over the First Two Billion Years. *ArXiv e-prints*, October 2014.

[15] K. Finlator, S. P. Oh, F. Özel, and R. Davé. Gas clumping in self-consistent reionization models. *MNRAS*, 427:2464–2479, December 2012.

[16] S. R. Furlanetto, M. Zaldarriaga, and L. Hernquist. The Growth of H II Regions During Reionization. *ApJ*, 613:1–15, September 2004.

[17] D. Galli and F. Palla. The Dawn of Chemistry. *ARAA*, 51:163–206, August 2013.

[18] E. M. George, C. L. Reichardt, K. A. Aird, B. A. Benson, L. E. Bleem, J. E. Carlstrom, C. L. Chang, H.-M. Cho, T. M. Crawford, A. T. Crites, T. de Haan, M. A. Dobbs, J. Dudley, N. W. Halverson, N. L. Harrington, G. P. Holder, W. L. Holzapfel, Z. Hou, J. D. Hrubes, R. Keisler, L. Knox, A. T. Lee, E. M. Leitch, M. Lueker, D. Luong-Van, J. J. McMahon, J. Mehl, S. S. Meyer, M. Millea, L. M. Mocanu, J. J. Mohr, T. E. Montroy, S. Padin, T. Plagge, C. Pryke, J. E. Ruhl, K. K. Schaffer, L. Shaw, E. Shirokoff, H. G. Spieler, Z. Staniszewski, A. A. Stark, K. T. Story, A. van Engelen, K. Vanderlinde, J. D. Vieira, R. Williamson, and O. Zahn. A Measurement of Secondary Cosmic Microwave Background Anisotropies from the 2500 Square-degree SPT-SZ Survey. *ApJ*, 799:177, February 2015.

[19] N. Y. Gnedin and A. H. Jaffe. Secondary Cosmic Microwave Background Anisotropies from Cosmological Reionization. *ApJ*, 551:3–14, April 2001.

[20] N. Y. Gnedin and J. P. Ostriker. Reionization of the Universe and the Early Production of Metals. *ApJ*, 486:581–598, September 1997.

[21] A. Gruzinov and W. Hu. Secondary Cosmic Microwave Background Anisotropies in a Universe Reionized in Patches. *ApJ*, 508:435–439, December 1998.

[22] O. Hahn and T. Abel. Multi-scale initial conditions for cosmological simulations. *MNRAS*, 415:2101–2121, August 2011.

[23] Z. Haiman, T. Abel, and P. Madau. Photon Consumption in Minihalos during Cosmological Reionization. *ApJ*, 551:599–607, April 2001.

[24] J. Harnois-Déraps, U.-L. Pen, I. T. Iliev, H. Merz, J. D. Emberson, and V. Desjacques. High-performance P^3M N-body code: $CUBEP^3M$. *MNRAS*, 436:540–559, November 2013.

[25] S. Hirano, T. Hosokawa, N. Yoshida, K. Omukai, and H. W. Yorke. Primordial star formation under the influence of far ultraviolet radiation: 1540 cosmological haloes and the stellar mass distribution. *MNRAS*, 448:568–587, March 2015.

[26] W. Hu. Reionization Revisited: Secondary Cosmic Microwave Background Anisotropies and Polarization. *ApJ*, 529:12–25, January 2000.

[27] I. T. Iliev, G. Mellema, K. Ahn, P. R. Shapiro, Y. Mao, and U.-L. Pen. Simulating cosmic reionization: how large a volume is large enough? *MNRAS*, 439:725–743, March 2014.

[28] I. T. Iliev, G. Mellema, U.-L. Pen, J. R. Bond, and P. R. Shapiro. Current models of the observable consequences of cosmic reionization and their detectability. *MNRAS*, 384:863–874, March 2008.

[29] I. T. Iliev, G. Mellema, P. R. Shapiro, and U.-L. Pen. Self-regulated reionization. *MNRAS*, 376:534–548, April 2007.

[30] I. T. Iliev, G. Mellema, P. R. Shapiro, U.-L. Pen, Y. Mao, J. Koda, and K. Ahn. Can 21-cm observations discriminate between high-mass and

low-mass galaxies as reionization sources? *MNRAS*, 423:2222–2253, July 2012.

[31] I. T. Iliev, U.-L. Pen, J. R. Bond, G. Mellema, and P. R. Shapiro. The Kinetic Sunyaev-Zel'dovich Effect from Radiative Transfer Simulations of Patchy Reionization. *ApJ*, 660:933–944, May 2007.

[32] I. T. Iliev, E. Scannapieco, and P. R. Shapiro. The Impact of Small-Scale Structure on Cosmological Ionization Fronts and Reionization. *ApJ*, 624:491–504, May 2005.

[33] I. T. Iliev and P. R. Shapiro. The post-collapse equilibrium structure of cosmological haloes in a low-density universe. *MNRAS*, 325:468–482, August 2001.

[34] I. T. Iliev, P. R. Shapiro, and A. C. Raga. Minihalo photoevaporation during cosmic reionization: evaporation times and photon consumption rates. *MNRAS*, 361:405–414, August 2005.

[35] A. H. Jaffe and M. Kamionkowski. Calculation of the Ostriker-Vishniac effect in cold dark matter models. *Phys. Rev. D*, 58(4):043001, August 1998.

[36] D. Jeong and E. Komatsu. Perturbation Theory Reloaded: Analytical Calculation of Nonlinearity in Baryonic Oscillations in the Real-Space Matter Power Spectrum. *ApJ*, 651:619–626, November 2006.

[37] A. A. Kaurov and N. Y. Gnedin. Cosmic Reionization On Computers III. The Clumping Factor. *ArXiv e-prints*, 1412.5607, December 2014.

[38] E. Komatsu, J. Dunkley, M. R. Nolta, C. L. Bennett, B. Gold, G. Hinshaw, N. Jarosik, D. Larson, M. Limon, L. Page, D. N. Spergel, M. Halpern, R. S. Hill, A. Kogut, S. S. Meyer, G. S. Tucker, J. L. Weiland, E. Wollack, and E. L. Wright. Five-Year Wilkinson Microwave Anisotropy Probe Observations: Cosmological Interpretation. *ApJS*, 180:330–376, February 2009.

[39] A. Lewis, A. Challinor, and A. Lasenby. Efficient Computation of Cosmic Microwave Background Anisotropies in Closed Friedmann-Robertson-Walker Models. *ApJ*, 538:473–476, August 2000.

[40] C.-P. Ma and J. N. Fry. Nonlinear Kinetic Sunyaev-Zeldovich Effect. *Physical Review Letters*, 88(21):211301, May 2002.

[41] P. Madau, F. Haardt, and M. J. Rees. Radiative Transfer in a Clumpy Universe. III. The Nature of Cosmological Ionizing Sources. *ApJ*, 514:648–659, April 1999.

[42] M. McQuinn, S. R. Furlanetto, L. Hernquist, O. Zahn, and M. Zaldarriaga. The Kinetic Sunyaev-Zel'dovich Effect from Reionization. *ApJ*, 630:643–656, September 2005.

[43] G. Mellema, I. T. Iliev, M. A. Alvarez, and P. R. Shapiro. C^2-ray: A new

method for photon-conserving transport of ionizing radiation. *New Astronomy*, 11:374–395, March 2006.

[44] L. Mercolli and E. Pajer. On the velocity in the Effective Field Theory of Large Scale Structures. *JCAP*, 3:6, March 2014.

[45] A. Mesinger, M. McQuinn, and D. N. Spergel. The kinetic Sunyaev-Zel'dovich signal from inhomogeneous reionization: a parameter space study. *MNRAS*, 422:1403–1417, May 2012.

[46] J. Miralda-Escudé, M. Haehnelt, and M. J. Rees. Reionization of the Inhomogeneous Universe. *ApJ*, 530:1–16, February 2000.

[47] P. A. Oesch, R. J. Bouwens, G. D. Illingworth, M. Franx, S. M. Ammons, P. G. van Dokkum, M. Trenti, and I. Labbe. First Frontier Field Constraints on the Cosmic Star-Formation Rate Density at z˜10 - The Impact of Lensing Shear on Completeness of High-Redshift Galaxy Samples. *ArXiv e-prints*, 1409.1228, September 2014.

[48] J. P. Ostriker and E. T. Vishniac. Generation of microwave background fluctuations from nonlinear perturbations at the ERA of galaxy formation. *ApJL*, 306:L51–L54, July 1986.

[49] C. Park. Cosmic momentum field and mass fluctuation power spectrum. *MNRAS*, 319:573–582, December 2000.

[50] H. Park, E. Komatsu, P. R. Shapiro, J. Koda, and Y. Mao. The Impact of Nonlinear Structure Formation on the Power Spectrum of Transverse Mo-

mentum Fluctuations and the Kinetic Sunyaev-Zel'dovich Effect. *ArXiv e-prints*, June 2015.

[51] H. Park, P. R. Shapiro, E. Komatsu, I. T. Iliev, K. Ahn, and G. Mellema. The Kinetic Sunyaev-Zel'dovich Effect as a Probe of the Physics of Cosmic Reionization: The Effect of Self-regulated Reionization. *ApJ*, 769:93, June 2013.

[52] A. H. Pawlik, J. Schaye, and E. van Scherpenzeel. Keeping the Universe ionized: photoheating and the clumping factor of the high-redshift intergalactic medium. *MNRAS*, 394:1812–1824, April 2009.

[53] J. A. Peacock and S. J. Dodds. Non-linear evolution of cosmological power spectra. *MNRAS*, 280:L19–L26, June 1996.

[54] S. Pueblas and R. Scoccimarro. Generation of vorticity and velocity dispersion by orbit crossing. *Phys. Rev. D*, 80(4):043504, August 2009.

[55] M. Raičević and T. Theuns. Modelling recombinations during cosmological reionization. *MNRAS*, 412:L16–L19, March 2011.

[56] C. L. Reichardt, L. Shaw, O. Zahn, K. A. Aird, B. A. Benson, L. E. Bleem, J. E. Carlstrom, C. L. Chang, H. M. Cho, T. M. Crawford, A. T. Crites, T. de Haan, M. A. Dobbs, J. Dudley, E. M. George, N. W. Halverson, G. P. Holder, W. L. Holzapfel, S. Hoover, Z. Hou, J. D. Hrubes, M. Joy, R. Keisler, L. Knox, A. T. Lee, E. M. Leitch, M. Lueker, D. Luong-Van,

J. J. McMahon, J. Mehl, S. S. Meyer, M. Millea, J. J. Mohr, T. E. Montroy, T. Natoli, S. Padin, T. Plagge, C. Pryke, J. E. Ruhl, K. K. Schaffer, E. Shirokoff, H. G. Spieler, Z. Staniszewski, A. A. Stark, K. Story, A. van Engelen, K. Vanderlinde, J. D. Vieira, and R. Williamson. A Measurement of Secondary Cosmic Microwave Background Anisotropies with Two Years of South Pole Telescope Observations. *ApJ*, 755:70, August 2012.

[57] B. E. Robertson, R. S. Ellis, J. S. Dunlop, R. J. McLure, and D. P. Stark. Early star-forming galaxies and the reionization of the Universe. *Nature*, 468:49–55, November 2010.

[58] R. Salvaterra, B. Ciardi, A. Ferrara, and C. Baccigalupi. Reionization history from coupled cosmic microwave background/21-cm line data. *MNRAS*, 360:1063–1068, July 2005.

[59] M. G. Santos, A. Cooray, Z. Haiman, L. Knox, and C.-P. Ma. Small-Scale Cosmic Microwave Background Temperature and Polarization Anisotropies Due to Patchy Reionization. *ApJ*, 598:756–766, December 2003.

[60] P. R. Shapiro, M. L. Giroux, and A. Babul. Reionization in a cold dark matter universe: The feedback of galaxy formation on the intergalactic medium. *ApJ*, 427:25–50, May 1994.

[61] P. R. Shapiro, I. T. Iliev, G. Mellema, K. Ahn, Y. Mao, M. Friedrich, K. Datta, H. Park, E. Komatsu, E. Fernandez, J. Koda, M. Bovill, and

U.-L. Pen. Simulating cosmic reionization and the radiation backgrounds from the epoch of reionization. In M. Umemura and K. Omukai, editors, *American Institute of Physics Conference Series*, volume 1480 of *American Institute of Physics Conference Series*, pages 248–260, September 2012.

[62] P. R. Shapiro, I. T. Iliev, and A. C. Raga. A model for the post-collapse equilibrium of cosmological structure: truncated isothermal spheres from top-hat density perturbations. *MNRAS*, 307:203–224, July 1999.

[63] P. R. Shapiro, I. T. Iliev, and A. C. Raga. Photoevaporation of cosmological minihaloes during reionization. *MNRAS*, 348:753–782, March 2004.

[64] L. D. Shaw, D. H. Rudd, and D. Nagai. Deconstructing the Kinetic SZ Power Spectrum. *ApJ*, 756:15, September 2012.

[65] J. M. Shull, A. Harness, M. Trenti, and B. D. Smith. Critical Star Formation Rates for Reionization: Full Reionization Occurs at Redshift $z \approx 7$. *ApJ*, 747:100, March 2012.

[66] R. E. Smith, J. A. Peacock, A. Jenkins, S. D. M. White, C. S. Frenk, F. R. Pearce, P. A. Thomas, G. Efstathiou, and H. M. P. Couchman. Stable clustering, the halo model and non-linear cosmological power spectra. *MNRAS*, 341:1311–1332, June 2003.

[67] G. C. So, M. L. Norman, D. R. Reynolds, and J. H. Wise. Fully Coupled Simulation of Cosmic Reionization. II. Recombinations, Clumping Factors, and the Photon Budget for Reionization. *ApJ*, 789:149, July 2014.

[68] V. Springel. The cosmological simulation code GADGET-2. *MNRAS*, 364:1105–1134, December 2005.

[69] V. Springel, M. White, and L. Hernquist. Erratum: Hydrodynamic Simulations of the Sunyaev-Zeldovich Effect(s). *ApJ*, 562:1086–1086, December 2001.

[70] V. Springel, S. D. M. White, G. Tormen, and G. Kauffmann. Populating a cluster of galaxies - I. Results at z=0. *MNRAS*, 328:726–750, December 2001.

[71] H. Trac and R. Cen. Radiative Transfer Simulations of Cosmic Reionization. I. Methodology and Initial Results. *ApJ*, 671:1–13, December 2007.

[72] D. Tseliakhovich and C. Hirata. Relative velocity of dark matter and baryonic fluids and the formation of the first structures. *Phys. Rev. D*, 82(8):083520, October 2010.

[73] E. T. Vishniac. Reionization and small-scale fluctuations in the microwave background. *ApJ*, 322:597–604, November 1987.

[74] S. Weinberg. *Cosmology*. Oxford University Press, Oxford, UK, 2008.

[75] N. Yoshida, S. P. Oh, T. Kitayama, and L. Hernquist. Early Cosmological H II/He III Regions and Their Impact on Second-Generation Star Formation. *ApJ*, 663:687–707, July 2007.

[76] N. Yoshida, K. Omukai, L. Hernquist, and T. Abel. Formation of Primordial Stars in a ΛCDM Universe. *ApJ*, 652:6–25, November 2006.

[77] O. Zahn, C. L. Reichardt, L. Shaw, A. Lidz, K. A. Aird, B. A. Benson, L. E. Bleem, J. E. Carlstrom, C. L. Chang, H. M. Cho, T. M. Crawford, A. T. Crites, T. de Haan, M. A. Dobbs, O. Doré, J. Dudley, E. M. George, N. W. Halverson, G. P. Holder, W. L. Holzapfel, S. Hoover, Z. Hou, J. D. Hrubes, M. Joy, R. Keisler, L. Knox, A. T. Lee, E. M. Leitch, M. Lueker, D. Luong-Van, J. J. McMahon, J. Mehl, S. S. Meyer, M. Millea, J. J. Mohr, T. E. Montroy, T. Natoli, S. Padin, T. Plagge, C. Pryke, J. E. Ruhl, K. K. Schaffer, E. Shirokoff, H. G. Spieler, Z. Staniszewski, A. A. Stark, K. Story, A. van Engelen, K. Vanderlinde, J. D. Vieira, and R. Williamson. Cosmic Microwave Background Constraints on the Duration and Timing of Reionization from the South Pole Telescope. *ApJ*, 756:65, September 2012.

[78] O. Zahn, M. Zaldarriaga, L. Hernquist, and M. McQuinn. The Influence of Nonuniform Reionization on the CMB. *ApJ*, 630:657–666, September 2005.

[79] P. Zhang, U.-L. Pen, and H. Trac. Precision era of the kinetic Sunyaev-

Zel'dovich effect: simulations, analytical models and observations and the power to constrain reionization. *MNRAS*, 347:1224–1233, February 2004.

[80] P. Zhang, Y. Zheng, and Y. Jing. Sampling artifact in volume weighted velocity measurement. I. Theoretical modeling. *Phys. Rev. D*, 91(4):043522, February 2015.